Prostate cancer stem cells: potential new biomarkers

Sharpe, Benjamin

Award date:
2016

Awarding institution:
University of Bath

PROSTATE CANCER STEM CELLS: POTENTIAL NEW BIOMARKERS

Benjamin Peter Sharpe

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

September 2016

ABSTRACT

Prostate cancer is a leading cause of cancer-related death in men, and while many men diagnosed with the disease will have an indolent clinical course, 20-25% of men will experience disease recurrence which is invariably lethal. There is an urgent need for prognostic biomarkers that will predict disease recurrence and risk-stratify patients upon diagnosis, allowing for personalised therapies. This thesis attempts to identify new prognostic biomarkers for prostate cancer and investigates their patterns of protein expression in human primary prostate tumour tissue.

Cancer stem cells are cancer cells thought to be uniquely capable of self-renewal and tumorigenicity, and may have a role in tumour recurrence. Using a literature searching approach, potential biomarkers related to stem cells, cancer stem cells or recurrence in prostate cancer were identified, and ALDH7A1, BMI1, SDC1, MUC1-C, Nestin and ZSCAN4 were chosen for investigation. An *in silico* approach was also used for biomarker identification, with RS1 and SLC31A1 selected as their mRNA was found to be upregulated in recurrent tumours. The expression patterns of all 7 potential biomarkers were examined by immunohistochemistry on prostate tumour tissue and benign tissue from prostate biopsies and prostatectomies.

BMI1, ALDH7A1, MUC1-C and Nestin showed no relationship to recurrence or other clinical features. RS1 protein levels increased in patients with recurrence within 5 years, negatively correlated with AR expression, and a meta-analysis showed that the RS1 gene was amplified in up to 32% of castration-resistant prostate tumours. ZSCAN4 was heterogeneously expressed in a subset of 26% of prostate tumours with unclear characteristics and was not expressed in benign tissue, but was not associated with recurrence. Finally, SDC1 expression was lost in tumour epithelium, but a population of unidentified SDC1-expressing cells were found in the stroma of a third of tumours, and an increased burden of these cells was associated with primary Gleason pattern 5 tumours. These cells do not overlap with common epithelial, mesenchymal or stromal lineages, but may be migratory. In summary, the data presented in this thesis identifies 3 potential new biomarkers for prostate cancer, and provides the basis for future characterisation of their wider roles in the disease.

ACKNOWLEDGEMENTS

Four years doesn't seem like a long time at all. There were times when I didn't think I could get this far, despite all evidence pointing to the contrary. I am writing this now in no small amount of gratitude to those people in my life who have encouraged me, brought out the best parts of me, and given me the strength to see myself through.

First and foremost, I am deeply grateful to my first supervisor, Dr Andrew Chalmers, for his support of me during my project, through the high points as much as the lows. We've known each other for about seven years now since you mentored me as a fresh-faced undergraduate, and your tireless optimism and boundless capacity for kindness never ceases to amaze me. I've learned so much, and I hope that the experience has been just as enjoyable and illuminating for you as it has for me. Rebecca, Mark and John: thank you for putting up with my constant requests for more samples and waiting patiently while I describe bioinformatics and what the blobs of red, green and blue are.

I would like to thank past and present members of the Chalmers, Wood, Williams, Caunt and Licchesi labs for sharing the office with me and providing excellent company and banter. In no particular order: Tori, Poonam, Gail, Carla, Bernard, Chris, Sophie, Kate, Kiri, Fred, Natalie, Dhafer, Cassidy and Tülay. I owe you all individual thanks. You had confidence in me from the start and it never faltered, even when my own did. I hope you've enjoyed at least some of my ramblings, scientific or otherwise.

I would like to thank Adrian Rogers for lending me tiny pieces of his vast imaging expertise. My gratitude also goes to Prof. Robert Kelsh and the members of his lab for allowing me to use their laboratory equipment, and especially to Karen and to Kleio for their company and conversation. Thank you both for sharing with me the joys of seasoning fruit salad with chilli, and zebrafish named Steve.

I would also like to thank Hazel and Jen, my brilliant friends with amazing minds and even more amazing hugs. Brady and Cole, your support has been unwavering and provided warmth and kindness no matter the circumstances. You have all given me the strength to do things I never would have thought myself capable of.

Finally, I owe a great deal of thanks to my family, who have supported me in so many ways from the very start. You provided immense strength during my lows and listened patiently to my successes. To my parents and my sister, who have given me patient, kind and unconditional support, guidance and love: I can never thank you enough.

CONTENTS

List of Tables

List of Figures

List of Abbreviations and Acronyms

ABCA3	ATP Binding Cassette Subfamily A Member 3
ABCG2	ATP Binding Cassette Subfamily G member 2
AFS	Anterior Fibromuscular Stroma
ALDH	Aldehyde Dehydrogenase
ALDH7A1	Aldehyde Dehydrogenase 7 Member A1
α2β1	α2β1 integrin
AMACR	Alpha-Methylacyl-CoA Racemase
APTS	(3-Aminopropyl)triethoxysilane
AR	Androgen Receptor
AURKA	Aurora Kinase A
BCL-2	B-Cell Lymphoma 2
BMI1	BMI1 Polycomb Ring Finger Oncogene
BPH	Benign Prostatic Hyperplasia
BrdU	Bromodeoxyuridine
C-myc	Myelocytomatosis Viral Oncogene Homolog
CARN	Castration-Resistant NKX3-1-Expresssing Luminal Cell
CDK2	Cyclin-Dependent Kinase 2
ChIP-Seq	Chromatin immunoprecipitation and sequencing
CRPC	Castration-Resistant Prostate Cancer
CSC	Cancer Stem Cell
CZ	Central Zone
DAPI	4',6-diamidino-2-phenylindole
DCLK1	Doublecortin-Like Kinase
DNA	Deoxyribonucleic Acid
E2F1	E2F Transcription Factor 1

ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic Acid
EGF	Epidermal Growth Factor
EHF	ETS Homologous Factor
eIF4 γ	Eukaryotic Translation Initiation Factor 4 Gamma
eIF5E	Eukaryotic Translation Initiation Factor 5 E
ELISA	Enzyme-linked Immunosorbent Assay
EMT	Epithelial-Mesenchymal Transition
ERG	V-Ets Avian Erythroblastosis Virus E26 Oncogene Homolog
ERK	Extracellular Signal-Regulated Kinase
ETS	ETS Proto-Oncogene
ETV	ETS Variant
EZH2	Enhancer of Zeste Homolog 2
FDR	False-Discovery Rate
FFPE	Formalin-Fixed Paraffin-Embedded
FGF	Fibroblast Growth Factor
FISH	Fluorescent In Situ Hybridisation
FLI1	Friend Leukaemia Virus Integration 1
FOXP3	Forkhead Box P3
GEO	Gene Expression Omnibus
GFP	Green Fluorescent Protein
GO	Gene Ontology
GSTP1	Glutathione S-Transferase Pi 1
H&E	Haematoxylin and Eosin
HCl	Hydrochloric acid
HER2	Human Epidermal Growth Factor Receptor 2

HGF	Hepatocyte Growth Factor
HIER	Heat-Induced Epitope Retrieval
HMGA2	High Mobility Group AT-Hook 2
HMWCK	High Molecular Weight Cytokeratin
hnRNPA2	Heterogeneous nuclear Ribonucleoprotein A2
HOOK3	Hook Microtubule-Tethering Protein 3
IF	Immunofluorescence
IGF1	Insulin-like Growth Factor 1
IGF1R	Insulin-like Growth Factor 1 Receptor
IGFBP-3	Insulin-like Growth Factor Binding Protein 3
IHC	Immunohistochemistry
Ikkα	I-Kappa-B Kinase Alpha
IMPDH2	Inosine Monophosphate Dehydrogenase 2
IRAS	Integrated Research Application System
K14	Cytokeratin 14
K18	Cytokeratin 18
K5	Cytokeratin 5
K8	Cytokeratin 8
KLF4	Kruppel-like Factor 4
LHRH	Leuteinising Hormone Releasing Hormone
Lin	Lineage. Cells expressing mature cell lineage markers.
mRNA	Messenger Ribonucleic Acid
MSI-1	Musashi RNA Binding Protein 1
mtDNA	Mitochondrial Deoxyribonucleic Acid
MUC1-C	Mucin 1, C-terminal Domain
NEPC	Neuroendocrine Prostate Cancer

NF- κB	Nuclear Factor Kappa B
NH3	Ammonia
NHS	National Health Service
NKX3-1	NK3 Homeobox 1
NOD/SCID	Non-Obese Diabetic/Severe Combined Immunodeficient mice
NRBP1	Nuclear Receptor Binding Protein 1
OCT3/4	Octamer-Binding Protein 4
ORF	Open Reading Frame
RNA-Seq	Ribonucleic acid sequencing
PBS	Phosphate-Buffered Saline
PCa	Prostate Cancer
PCA3	Prostate Cancer Associated 3
PCNA	Proliferative Cell Nuclear Antigen
PCSP	Prostate Cancer Syndecan-1 Positive Stromal Cell
PD-1	Programmed Cell Death 1
PD-L1	Programmed Cell Death Ligand 1
PDGF	Platelet-Derived Growth Factor
PHLPP	PH Domain And Leucine Rich Repeat Protein Phosphatase
PI3K	Phosphoinositide 3-Kinase
PIA	Proliferative Inflammatory Atrophy
PIN	Prostatic Intraepithelial Neoplasia
PIP3	Phosphatidylinositol triphosphate
PRC	Polycomb Repressor Complex
PSA	Prostate-Specific Antigen
PSAP	Prostatic Acid Phosphatase
PTEN	Phosphatase And Tensin Homologue

PZ	Peripheral Zone
qRT-PCR	Quantitative Reverse Transcriptase Polymerase Chain Reaction
RAS	Rat Sarcoma Viral Oncogene Homolog
REC	Research Ethics Committee
RhoA	Ras Homolog Family Member A
RNA	Ribonucleic Acid
RS1	Retinoschisin 1
SDC1	Syndecan-1
SILAC	Stable Isotope Labelling with Amino Acids in Cell Culture
siRNA	Small interfering Ribonucleic Acid
SLC31A1	Solute Carrier Family 31 Member 1
SLC45A3	Solute Carrier Family 45 Member 3
SMO	Smoothened, Frizzled Class Receptor
SOX2	SRY-Box 2
SPINK1	Serine Peptidase Inhibitor, Kazal Type 1
SPOP	Speckle Type BTB/POZ Protein
SRC	Steroid Receptor Coactivator
SV	Seminal Vesicle
TCF4	Transcription Factor 4
TCGA	The Cancer Genome Atlas
TMA	Tissue Microarray
TMPRSS2	Transmembrane Protease, Serine 2
TRIM24	Tripartite Motif Containing 24
TURP	Transurethral Resection of the Prostate
TZ	Transition Zone
Max	Myc Associated Factor X

WOX1	WW Domain Containing Oxidoreductase 1
ZSCAN4	Zinc Finger and SCAN Domain Containing 4

List of Appendices

1 INTRODUCTION

Prostate cancer is a global healthcare burden, and a heterogeneous disease which presents with unique challenges to effective diagnosis and treatment. One of the largest challenges in the field is to identify features that will allow patients to be risk-stratified, tailoring their treatments to their individual needs. This thesis will attempt to address this challenge by identifying new biomarkers to predict recurrence and risk-stratify patients with prostate cancer.

1.1 The Prostate

The following section describes the prostate, its normal biological functions, anatomy and histology, and discusses the ongoing research with regards to understanding of the cell types and lineage hierarchies in the developing and adult prostate. This provides the basis for understanding the molecular processes involved in prostate cancer formation and progression which are discussed in later sections.

1.1.1 Physiology and anatomy of the prostate

1.1.1.1 Prostate physiology

The prostate is a glandular organ located beneath the bladder and anterior to the rectum. Its main physiological function is the secretion of prostatic fluid, which contributes to approximately 30% by volume of the seminal fluid and promotes sperm motility and survival in the seminal plasma (Burden *et al.*, 2006; Ousset *et al.*, 2012). Chief components of the prostatic fluid include citrate and zinc, the levels of which are particularly high and correlate with one another in benign prostate cells (Costello and

Franklin, 1998). Zinc is chelated by citrate and has antibacterial properties (Burden *et al.*, 2006), while citrate serves as an energy source for spermatozoa (Medrano *et al.*, 2006). It is also the origin of the serine protease Prostate-Specific Antigen (PSA) which is responsible for cleaving seminal semenogelins and fibronectin in order to liquefy the seminal plasma clot and promote sperm motility in the female reproductive tract (Yousef and Diamandis, 2001; Lilja *et al.*, 2008).

The prostate is an androgen-dependent organ, being regulated by exposure to androgens synthesised in the testes and in the adrenal glands through expression of the Androgen Receptor (AR). Secretory functions of the prostate are regulated in an androgen-responsive manner, a phenomenon demonstrated in early canine experiments by measuring secretion volumes following androgen and estrogen administration (Huggins and Sommer, 1953) and importantly many prostate epithelial cells require maintained levels of androgen and AR signalling to survive (Davis and Day, 2002; Kurita *et al.*, 2001). Early murine experiments showed that castration of adult rats results in apoptosis of prostate epithelial cells and subsequent involution of the gland (Kyprianou and Isaacs, 1988) – a finding replicated many times experimentally in mice and used as a model of adult prostate homeostasis and stem cell activity (Choi *et al.*, 2012; Wang *et al.*, 2009; Wang *et al.*, 2013). Androgen-dependent epithelial cell survival is dependent on stromal androgen receptor activity, and appears to be a paracrine signalling effect of the stroma on epithelial cells, as even AR-negative prostate epithelium will undergo apoptosis if androgen is withdrawn in the context of an AR-expressing stromal microenvironment (Kurita *et al.*, 2001).

1.1.1.2 Prostate anatomy

Anatomically, the prostate is a branched ductal structure which surrounds the urethra. Ducts are embedded in a fibromuscular stroma such that contractions empty the contents of the ductal lumens into the urethra (Fine and Reuter, 2012). According to publications by McNeal in the 1980s, who drew anatomical conclusions from dissecting autopsy specimens, this structure is divided into three glandular zones: the transition zone surrounding the urethra at the mid-gland region, the central zone surrounding the ejaculatory ducts, and the peripheral zone surrounding and encompassing these two regions to make up the outer portion of the prostate (McNeal, 1981) (Figure 1.1). The prostate also has an anterior fibromuscular zone sitting anteriorly to the urethra and transition zones (McNeal, 1981). Surrounding the prostate is a layer of fibromuscular

tissue known as the prostatic capsule (Fine and Reuter, 2012). The peripheral zone is the largest of the glandular zones, accounting for approximately 65% of the prostate and is the most common site of occurrence for prostate carcinoma (Fine and Reuter, 2012). The central zone accounts for 30% of the prostate mass and the transition zone only 5% (Fine and Reuter, 2012; McNeal *et al.*, 1988). The transition zone is the site at which a benign hyperplastic condition unrelated to carcinoma arises, known as benign prostatic hyperplasia (BPH), but the region varies considerably in size and positioning between individuals (Fine and Reuter, 2012). Carcinoma in this region is less common and usually assumes a distinct 'clear cell' histological appearance (McNeal *et al.*, 1988; Fine and Reuter, 2012).

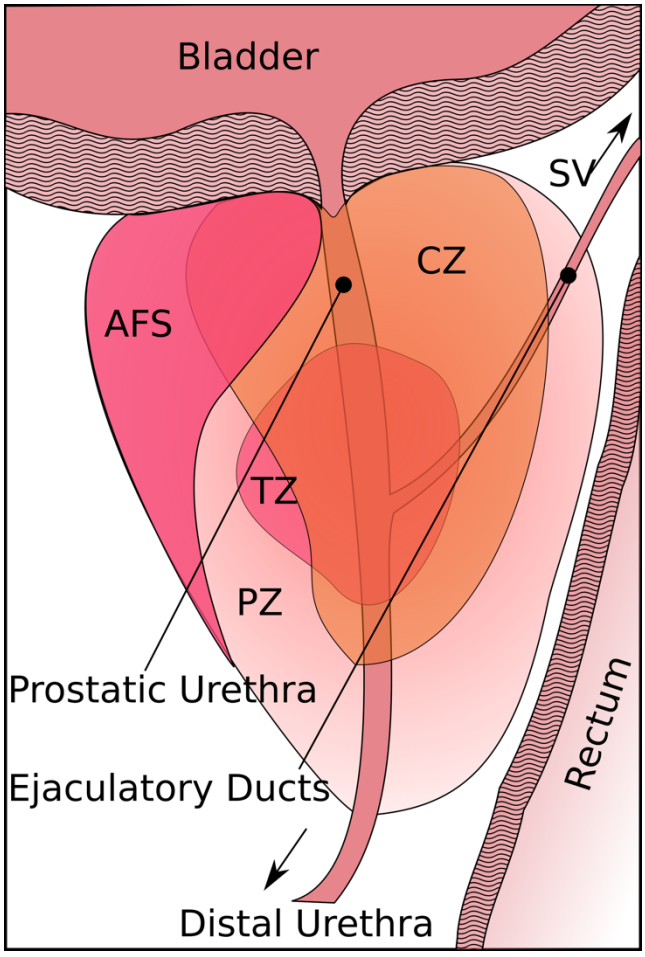


Figure 1.1: Diagram of prostate anatomy as described by McNeal.

The prostate encircles the urethra consists of three glandular zones: the transition zone (TZ), central zone (CZ) and peripheral zone (PZ). A region of anterior fibromuscular stroma (AFS) is also present. Seminal vesicles (SV) also secrete their fluid into the urethra through the ejaculatory ducts, labelled.

1.1.2 Prostate Histology

The prostate gland consists of an array of branched ductal structures with a pseudostratified columnar epithelial lining surrounded by a fibromuscular stroma. The prostate epithelium has three main cell types: basal cells which line the basement membrane (De Marzo *et al.*, 1999; McNeal, 1988); luminal cells, which are columnar cells that fulfil the secretory functions of the prostate and sit atop the basal layer (McNeal, 1988); and neuroendocrine cells – an "endocrine-paracrine" cell type which appears scattered infrequently through the prostate epithelium between basal and luminal cells (Taylor and Risbridger, 2008; McNeal, 1988), and secretes serotonin and regulatory peptide hormones (diSantAgnese and Cockett, 1996). Figure 1.2 shows a diagram of the histology of the normal prostate with all epithelial cell types labelled.

1.1.2.1.1 Basal cells

Basal cells are in contact with the basement membrane of the prostatic ducts, and in humans form a continuous layer with apparent cell-cell contacts (El-Alfy *et al.*, 2000) whereas in other mammals the layer is discontinuous and form a basal-to-luminal ratio of approximately 1:7. Basal cells express high molecular weight cytokeratins identified by the antibody clone 34BE12 which includes cytokeratins 5 and 14 (De Marzo *et al.*, 1999), and the p53 homologue p63 that is essential for the development of basal cells (Kurita *et al.*, 2004). It is hypothesised that the basal cell compartment contains putative prostate stem cells, although in humans it has been shown that they also carry ultramicroscopic indicators of being terminally differentiated, such as glycogen particle accumulation and a large Golgi apparatus (El-Alfy *et al.*, 2000). The function of basal cells is not fully understood, as they do not perform secretory roles and are AR- and PSA-negative (Taylor and Risbridger, 2008), but they do serve to completely separate the luminal cells from the stromal compartment which suggests a regulatory role for the cell type in humans (El-Alfy *et al.*, 2000; Kurita *et al.*, 2004), especially as basal cells communicate with luminal cells through gap junctions (El-Alfy *et al.*, 2000).

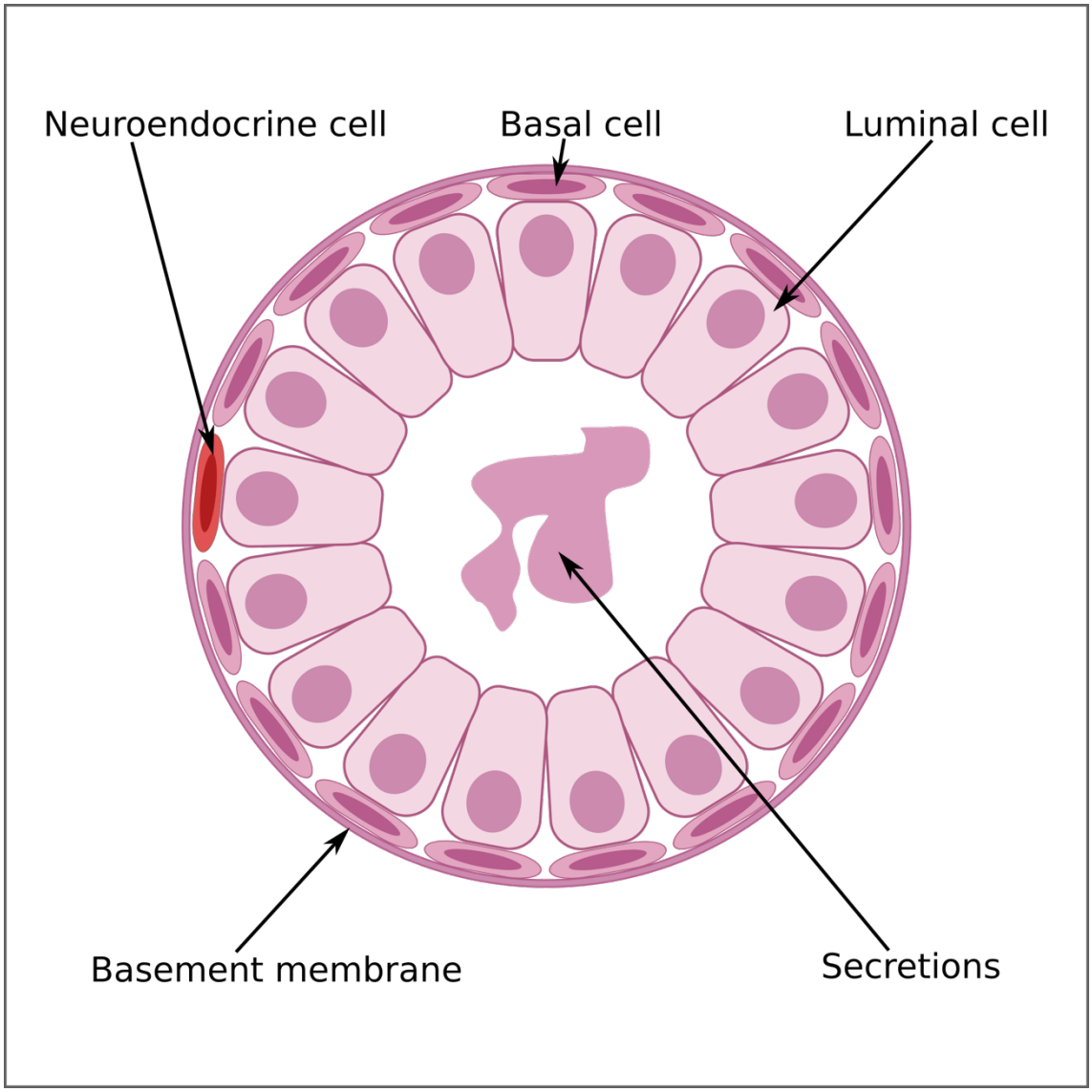


Figure 1.2: Histology of the normal prostate epithelium.

The normal prostate duct is an epithelial structure with a basement membrane, upon which a layer of flattened basal cells with minimal cytoplasm sit. Luminal cells sit upon the layer of basal cells at a 1:1 ratio and secrete substances such as zinc, citrate and PSA into the ductal lumen (labelled). In addition to these two main cell types, there is also a neuroendocrine cell type which secretes neuropeptides and is often found in the basal cell layer, histologically indistinguishable from the basal cells.

1.1.2.1.2 Luminal cells

Luminal cells express nuclear AR, PSA, and Prostate Specific Acid Phosphatase (PSAP), as well as secretory cytokeratins 8 and 18 as identified by the Cam 5.2 monoclonal antibody (De Marzo *et al.*, 1999). These serve the secretory functions of the prostate as previously mentioned in Chapter 1.1.1.1.

1.1.2.1.3 Neuroendocrine cells

Neuroendocrine prostate cells are defined by dendritic processes and the secretion of serotonin and a number of neuropeptides: among them are chromogranins, calcitonin family peptides, parathyroid-like hormone, and somatostatin (diSantAgnese and Cockett, 1996; Hirano *et al.*, 2004) as well as bombesin and vascular endothelial growth factors (Hirano *et al.*, 2004; Ismail *et al.*, 2002). They are commonly identified *in situ* by immunostaining for the markers chromogranin A (Mucci *et al.*, 2000; Hirano *et al.*, 2004; Ahlgren *et al.*, 2000) and synaptophysin (Mucci *et al.*, 2000) or serotonin (Ismail *et al.*, 2002). These neuropeptide secretory roles make them similar to neuroendocrine cells described in other organs, such as the intestine, lung, pancreas and thyroid (Ismail *et al.*, 2002; diSantAgnese and Cockett, 1996).

Neuroendocrine cells are innervated by efferent autonomic nerves and sensory afferent nerves (diSantAgnese and Cockett, 1996). Their presence is infrequent within the prostate epithelium and they are usually localised within the luminal cell layer, either exposed to the luminal space or surrounded by luminal cells instead (Ismail *et al.*, 2002; diSantAgnese and Cockett, 1996). Unlike luminal cells, neuroendocrine cells do not express the androgen receptor whether in a benign or cancerous context, and do not respond to hormone deprivation (Krijnen *et al.*, 1993). Interestingly, they seem to increase in frequency in human prostate cancer xenografts in mice when subjected to castration (Huss *et al.*, 2004) and also in patients with castration-resistant prostate tumours (Hirano *et al.*, 2004; Ahlgren *et al.*, 2000), suggesting a relationship between prostatic androgen exposure and neuroendocrine differentiation.

1.1.3 Adult prostate maintenance and prostate stem cells

In tissues with rapid renewal and turnover, such as the skin and colon, there exists a long-lived adult stem cell population which are thought to both self-renew and gives rise to transit-amplifying cells: cells with a limited self-renewal potential but the ability to divide asymmetrically to give rise to differentiated cell types (Taylor and Risbridger, 2008). A stem cell niche is thought to exist in the prostate, as it is capable of more than

fifteen rounds of serial androgen withdrawal and replacement, which causes most luminal and some basal epithelial cells to undergo apoptosis, before restoring normal glandular architecture and cell types upon androgen replacement (Wang *et al.*, 2009). However, tissue renewal and cell turnover is much slower in the prostate than in the skin and colon, and heterogeneity in the system makes elucidation of a defined hierarchy difficult. In addition, any model of the prostate stem cell hierarchy must be able to take into account all three lineages of prostate cells: basal cells, luminal cells and neuroendocrine cells.

Models have been hypothesised which include multipotent progenitors giving rise to multiple cell lineages, and/or unipotent progenitors that can only give maintain one cell type (Ousset *et al.*, 2012). A study involving BrdU incorporation found that both basal and luminal cells were more slowly dividing in proximal than distal regions of murine prostate ducts, and these cells possessed the capacity to efficiently generate glandular structures in vitro (Tsujimura *et al.*, 2002). In an attempt at 'in situ lineage tracing', mtDNA sequencing has been performed in cytochrome c oxidase-deficient and functioning areas of prostate epithelium, which accumulate mtDNA mutations over the course of years and decades, to predict the clonal origin of cells within each prostate acinus (Blackwood *et al.*, 2011). The study identifies regions of shared clonal origin within individual acini, which suggests there may be a multipotent stem cell to give rise to basal, luminal and neuroendocrine cells in the human prostate (Blackwood *et al.*, 2011). However, it is currently unknown whether such a cell resides in the luminal or basal compartments of the epithelium. A knowledge of the hierarchy of prostate development and differentiation has the potential to inform future models of prostate cancer initiation and progression and so will be discussed here.

In early experiments, proliferative markers such as Proliferative Cell Nuclear Antigen (PCNA) (McNeal *et al.*, 1995; Bonkhoff *et al.*, 1994) and Ki-67 (Feneley *et al.*, 1996; Bonkhoff *et al.*, 1994) were found to be present in a small proportion of basal cells in the benign human prostate epithelium. Mice carrying a homozygous deletion at the p63 locus, which is expressed exclusively by the basal epithelial cells within the prostate, are unable to develop the organ proper (Signoretti *et al.*, 2000). This suggests that the basal cell population contains a putative stem cell population responsible for prostate development and all the epithelial cell types found within. Van Leenders *et al.* demonstrated that a subset of basal prostate epithelial cells can express the canonical basal and luminal markers K5/K14/K18 simultaneously in human prostate tissue

sections and in culture (2000). They suggest that these correspond to stem cells whereas most basal cells, which are only K5/K14+, are a transit-amplifying or intermediate cell type that will lose K5/K14 expression and differentiate to become luminal cells with a K8/K18 expression profile. Although the group found cells with neurite-like processes suggestive of neuroendocrine differentiation in their explant culture system, they were unable to demonstrate neuroendocrine differentiation in the model (van Leenders *et al.*, 2000). Similarly, mice which bear a prostate tissue graft that is homozygous for a deletion at the p63 locus, which is expressed in prostate basal cells and necessary for murine basal cell development, develop synaptophysin-positive neuroendocrine cells at a similar frequency to wild-type tissue grafts, suggesting that neuroendocrine cells may not differentiate from a basal cell subtype in a developmental context (Kurita *et al.*, 2004).

More recently, the subject has been studied using lineage tracing models in the mouse. The typical strategy for lineage tracing requires an inducible genetic recombination event that permanently marks specific cell types and all their progeny. Lineage-specific promoter-driven expression of a Cre recombinase enzyme is established in a transgenic organism and 'floxed' alleles containing a fluorescent reporter are designed to be made transcriptionally active following successful recombination (Wang *et al.*, 2009; Wang *et al.*, 2013). This allows the fates and identities of these cells and their progeny to be traced over time. Using this approach, Ousset et al demonstrated considerable heterogeneity during murine postnatal prostate development, with K5- and K14-expressing basal cells able to give rise to both basal and luminal cells, although some traced cells formed clones of solely luminal cells (Ousset *et al.*, 2012). In this setting, cells with an intermediate phenotype co-expressing luminal and basal markers were associated with bipotent clones that gave rise to luminal cells, suggesting that the intermediate phenotype is associated with luminal differentiation. The same group later used multi-colour lineage tracing to focus on postnatal prostate development and combined this with statistical modelling of the labelled cell populations, and found that lineage-marked basal cells gave rise to labelled luminal cells more than would be expected by chance (Wuidart *et al.*, 2016). This further confirms that, during prostate development, basal cells are bipotent in nature.

In the adult tissue homeostasis setting the situation may be less clear. It has been shown in one lineage-tracing study that in a model of castration and regeneration, at least a subpopulation of basal and luminal cells are unipotent and can generate more of their

own lineage, with no evidence of one lineage contributing to the other (Choi *et al.*, 2012). A later lineage-tracing study suggests that a switch from a bipotent to unipotent mode of cell differentiation occurs at approximately postnatal day 30, when the percentage of lineage-marked luminal cells remains relatively constant (Wuidart *et al.*, 2016).

Other lineage-tracing studies show a less clear picture for the luminal cell lineage in the adult prostate. Another lineage-tracing study, which includes more rounds of castration and prostate regeneration, demonstrates that basal cells can give rise to luminal cells, although this is a rare phenomenon constituting 0.04% of lineage-marked basal cells (Wang *et al.*, 2013). Tracking lineage-marked mice over a course of a year without castration confirmed that differentiation of bipotent basal cells occurs at a low frequency during normal prostate homeostasis (Wang *et al.*, 2013). Similarly, a seminal study isolated a putative prostate stem cell population from mice which consisted of as few as 0.12% of prostate cells, and was identified by the marker panel Lin-/Sca1+/CD133+/CD44+/CD117+, and was able to generate all three prostate lineages and the appropriate prostate duct architecture from a single cell (Leong *et al.*, 2008). CD117 (c-kit)-expressing basal cells exist in human prostate tissue, but it has yet to be demonstrated that they have a similar stem-like potential (Leong *et al.*, 2008).

It has also been shown that a subset of NKX3-1 expressing luminal cells are resistant to castration and can regenerate all prostate epithelial cell types from a single cell, suggesting that there could be multipotent stem cells residing in the luminal compartment as well (Wang *et al.*, 2009). NKX3-1 appears to be critical for prostatic differentiation: lentiviral expression of NKX3-1 in seminal vesicle tissue in a renal graft is sufficient to convert it to a prostate-like glandular structure with apparent basal and luminal cell layers (Dutta *et al.*, 2016), and expression of NKX3-1 in the basal-like prostate cell line RWPE1 renders it similarly capable of generating secretory ductal prostate-like structures in renal grafts.

In summary, lineage-tracing experiments demonstrate evidence for both multipotent and unipotent cells in the basal and luminal epithelial cell compartments, but due to the experimental models employed, which largely rely on murine prostate regeneration and grafting assays, the picture of what occurs during normal prostate homeostasis is far from clear. In addition, it is worth noting that most stem cell characterising experiments

have been conducted in murine prostate cells, and the murine prostate is anatomically and structurally distinct from the human prostate (Abate-Shen and Shen, 2000).

1.2 Prostate cancer and other prostatic disease

Prostate adenocarcinoma is the most common form of prostate cancer and is a malignant hyperplastic disease arising from the epithelial compartment of the prostate gland. It is one of the most common cancers in men and a leading cause of cancer-related death, with the USA's Surveillance, Epidemiology and End Results Program (SEER) predicting over 180,000 new diagnoses and over 26,000 deaths of prostate cancer in this year in the USA alone (Siegel *et al.*, 2016). This accounts for 1 in 5 of all new cancer diagnoses in men, and an estimated 1 in 7 men will contract the disease at some point in their lives (Siegel *et al.*, 2016). The following section examines the process of diagnosis and therapy for prostate carcinoma, the molecular pathology of the prostate carcinoma during its progression, and considers the utility of biomarkers for diagnosis and prognostication of the disease.

1.2.1 The clinical course of prostate adenocarcinoma

1.2.1.1 Diagnosis

Patients with prostate carcinoma are normally identified through PSA screening (refer to Chapter 1.2.2) or presentation to primary care with lower urinary tract symptoms, such as a poor urinary stream, urinary urgency or hesitancy (Hamilton and Sharp, 2004). However, these symptoms are common in other benign conditions such as benign prostatic hyperplasia, which is also a common hyperplastic disease in men (Hamilton and Sharp, 2004), and lower urinary tract symptoms do not confirm a diagnosis or add additional diagnostic power to an investigation. Therefore, care must be taken to distinguish patients with benign prostatic hyperplasia from patients with carcinoma. The histological methods for this will be covered in detail in Chapter 1.2.2. National Institute for Health and Care Excellence (NICE) guidelines currently state that patients with lower urinary tract symptoms should undergo a rectal examination and a PSA test (NICE, 2015). Rectal examination for a palpable prostate tumour is a more robust indicator of prostate cancer than an abnormal PSA reading (Hamilton and Sharp, 2004; Thompson *et al.*, 2004) and either an abnormal PSA test or an adverse finding on rectal examination warrants taking core needle biopsies of prostate tissue for histological evaluation (NICE, 2015). A PSA test revealing a serum PSA concentration

of greater 4ng/mL is usually considered abnormal, although this threshold is often disputed and men with lower levels can develop prostate cancer (Thompson *et al.*, 2004). Core needle biopsies are normally taken with 18 gauge needles and at least 10 biopsy cores should be taken, guided by transrectal ultrasound (Heidenreich *et al.*, 2008). Histological evaluation will be discussed in more detail after considering treatment (see Chapter 1.2.2).

1.2.1.2 Therapeutic intervention

1.2.1.2.1 Primary therapies

When patients are diagnosed with prostate cancer, the primary tumour is treated with one of several treatment modalities. The most common treatment modality for primary prostate carcinoma is radical prostatectomy (Whiting *et al.*, 2016), where the prostate is removed surgically through well-defined methods dating back to 1947 (Mukouyama *et al.*, 2001). The second most popular choices are external beam radiotherapy and brachytherapy (Whiting *et al.*, 2016), where ionising radiation is delivered through either an external radiation source or surgically implanted radioisotope-containing materials, respectively. In either case, there are considerable impacts of primary therapy on patient quality of life, with commonly reported symptoms of treatment including erectile dysfunction, urinary and/or bowel incontinence and impotence as well as anxiety and depression (Whiting *et al.*, 2016; Kundu *et al.*, 2004). Treatment with hormone therapy is less commonly recommended as a primary therapy, but acts by lowering serum androgen levels and inhibiting the AR signalling axis in prostate cells, which is generally required for prostate epithelial cell survival (Kyprianou and Isaacs, 1988). Common methods for hormone therapy include the administration of Luteinising Hormone Releasing Hormone agonists, which block testicular androgen synthesis, along with non-steroidal AR antagonists such as bicalutamide, flutamide or enzalutamide (Heidenreich *et al.*, 2008). There may be a survival benefit for early administration of hormone therapy in patients with locally advanced prostate cancer (Smith *et al.*, 2005), though this is still under debate and requires more randomised controlled trials to be conclusive (Heidenreich *et al.*, 2008).

In terms of outcomes for clinically localised disease, prostatectomy may have benefit over radiotherapy by reducing prostate cancer-specific mortality rates (Sooriakumaran *et al.*, 2014; Lee *et al.*, 2015a; D'Amico *et al.*, 2002), although this benefit is most apparent in patients with lower risk disease (D'Amico *et al.*, 2002) and outcomes may

be more equivalent in high risk disease (D'Amico *et al.*, 2002; Boorjian *et al.*, 2011). Reporting and standard classifications of recurrence differ under treatment with prostatectomy and radiotherapy, which precludes meaningful comparison between the two treatment groups, but there are indications that the less stringent reporting of serum PSA increases in radiotherapy could lead to a delay in reporting and thus increased risk to patients (Lee *et al.*, 2015a). This is especially the case as the establishment of a PSA nadir, where serum PSA levels remain consistently at a minimum point, takes much longer to reach following radiotherapy than for prostatectomy (D'Amico *et al.*, 2002), and this is part of the framework on which recurrence is assessed.

Prostate carcinoma is biologically heterogeneous along with considerable heterogeneity in clinical outcome, and is one of the few forms of cancer for which 10-year recurrence-free survival is a useful measure of outcomes due to the potentially long durations that can pass between treatment and recurrence (Bianco *et al.*, 2003). Although primary therapy is usually of curative intent, approximately 20-25% of patients experience recurrence after primary therapy, either as assessed by rising serum PSA levels, local recurrence or presence of metastases (Epstein *et al.*, 1993; Bianco *et al.*, 2003). Due to the impacts of primary therapy and recurrence, it is important to classify patients based on estimated disease severity.

1.2.1.2.2 Secondary therapies

The choice of secondary therapy is much more variable and depends upon patient preferences as well as the site of recurrence. Salvage radiotherapy or prostatectomy may be recommended for patients who have experienced local failure following prostatectomy or radiotherapy respectively (Heidenreich *et al.*, 2008). For patients with systemic failure, where the tumour has spread to distant locations including skeletal metastases, treatment with hormone therapy is recommended. As with primary hormonal therapy, secondary hormonal therapy usually consists of an LHRH agonist, although anti-androgens may be added when this treatment begins to fail (Sternberg *et al.*, 2013). However, in patients with recurrent disease, hormone therapy invariably fails, usually within 19-24 months (Sternberg *et al.*, 2013), and in this case Docetaxel chemotherapy is recommended as the last line of treatment, and patients generally respond for 12-29 months (Heidenreich *et al.*, 2008). Castration-resistant disease, where cancer continues to progress clinically despite castrate levels of testosterone (Sternberg *et al.*, 2013), is invariably lethal (Sung and Cheung, 2014), which means that predictors

of this progression and characterisation of the molecular pathways involved are vital for treatment of advanced prostate cancer. For more information on the pathways and alternations thought to be involved, see Chapter 1.2.3.9.

1.2.2 Detection, diagnosis, and challenges

1.2.2.1 Prostate-specific Antigen (PSA) and the PSA era

Both benign prostate and prostate carcinoma express PSA as part of the normal secretory programme of luminal prostate epithelial cells (Catalona *et al.*, 1991; De Marzo *et al.*, 1999; Lilja *et al.*, 2008). Serum PSA levels in healthy men are approximately 0.6ng/mL in normal prostate tissue, and rise in some cases of prostate cancer to 4ng/mL or greater (Catalona *et al.*, 1991). This is suspected to be due to the structural damage and disorder inherent to prostate carcinoma, resulting in greater PSA release into the blood rather than overexpression of the protein in tumours proper (Lilja *et al.*, 2008). This finding is used as a tool in both diagnostic and treatment monitoring settings for prostate cancer but has limitations, as outlined below.

1.2.2.1.1 PSA testing for screening and diagnosis

Routine screening of men for prostate cancer using a serum PSA test caused prostate cancer diagnoses to surge during the late 1980s and early 1990s (Siegel *et al.*, 2016). Patients may suffer from prostate cancer but not have palpable disease on rectal examination, and a PSA test is invaluable for diagnosing these cases that would otherwise have been missed (Catalona *et al.*, 1991). PSA screening has had the positive effect of shifting down the stages at which patients with prostate cancer present, meaning that most patients now present with small nodules of prostate cancer detected by screening or biopsy that would not be detectable by palpation or imaging, known as stage T1c (Bianco *et al.*, 2003). Concomitantly, the frequency of patients presenting with advanced disease and high grade tumours has dropped dramatically in the PSA screening era (Penney *et al.*, 2013).

However, routine screening of men for prostate cancer using the PSA test is no longer advocated, as an estimated 23% to 42% of screen-detected tumours are over-diagnosed (Siegel *et al.*, 2016), meaning that they present with clinically insignificant disease and are likely to expose the men to unnecessary risk in the forms of therapeutic intervention as well as the associated anxieties of cancer diagnosis. Because of this change in procedure in the USA, the number of diagnoses of prostate cancer in the USA are

beginning to fall (Siegel *et al.*, 2016). While the general consensus in Europe is much the same with regards to screening of the general population, PSA screening has been found to be sufficiently sensitive, specific and cost-effective compared to other cancer screening programmes (Crawford and Abrahamsson, 2008). The benefits of early detection and reduced disease progression rates arguably outweigh the risks of over-diagnosis. However, from its inception, PSA testing as a first line screening approach has had an additional flaw: half of all patients with an elevated PSA level, that would be considered suspect for carcinoma, have benign prostatic hyperplasia or prostatitis instead (Catalona *et al.*, 1991). Additionally, some men may have prostate cancer even with normal serum PSA levels (Heidenreich *et al.*, 2008). Thus, there is the potential for both false positive and false negative findings. As more men with prostate cancer die of other causes than due to the disease, and many more early-stage cancers are being detected, new biomarkers will be required to provide more specific cancer diagnosis and prognostication.

1.2.2.1.2 PSA testing for monitoring treatment course

Following therapy, the efficacy of a treatment can be monitored by routinely measuring the levels of PSA in the serum (Lilja *et al.*, 2008). This manifests in different ways for different treatments. For prostatectomy, serum PSA levels are expected to fall rapidly after surgery, reducing to levels of less than 0.2ng/mL (Heidenreich *et al.*, 2008). For radiotherapy, a serum PSA level of greater than 2ng/mL above the nadir PSA is considered to be recurrence, otherwise known as biochemical failure or biochemical recurrence (Roach *et al.*, 2006). Unfortunately, there is no consensus definition of biochemical recurrence, although a standard has been suggested as serum PSA >0.4ng/mL followed by an increase at the next measurement (Stephenson *et al.*, 2006).

PSA kinetics such as the doubling time, or the time duration required for a doubling of serum PSA levels, are often used to monitor disease progression following therapy (Slovin *et al.*, 2005). Another such measure is PSA velocity, which also measures the rate at which serum PSA levels increase over time. Both higher baseline PSA levels after treatment and higher PSA velocity are independent predictors of shorter survival and an increased risk of bone metastases in non-metastatic patients treated with hormone therapy (Smith *et al.*, 2005). Many men in this study had an indolent clinical course despite having rising PSA, suggesting that the presence of a rising PSA level following therapy may not be as informative as the levels and kinetics. In this way,

biochemical recurrence does not always lead to clinical progression, and clinical endpoints are always the more certain measures of patient progression and survival. A shorter PSA doubling time was also significantly associated with shorter metastasis-free survival in a study of patients treated with prostatectomy (Antonarakis *et al.*, 2012). However, in some patients metastatic progression of prostate carcinoma can occur without an accompanying rise in PSA levels, especially in patients with atypical histological variants of carcinoma such as small cell carcinoma (Leibovici *et al.*, 2007).

In summary, while testing serum PSA levels should not be used as the sole method of monitoring patient disease in a post-treatment setting, it remains a highly useful tool for routinely monitoring the response of patients to their therapies, and acts as a decision-making tool guiding the timings of subsequent therapy selection.

1.2.2.2 The Gleason grading system and tumour staging

There are two systems in place in the clinic for categorising patients into risk categories. These are the Gleason grading system and tumour staging.

The Gleason pattern was originally developed by Dr. Donald Gleason in 1974 as a means of classifying prostate tumour morphology and associating it with clinical outcome, based on the observation that there are distinct patterns of tumour growth and that there tend to be two or more major growth patterns within a tumour biopsy (Gleason and Mellinge.Gt, 1974). The grading system has been altered over the years, but is still largely descriptive and represents the degree of morphological differentiation of the tumour by H&E staining on a scale of 1 to 5. Pattern 1 is the most well differentiated and 5 is poorly differentiated, with higher Gleason patterns assigned to increasingly small, irregular or atypical glandular structures (Gordetsky and Epstein, 2016; Humphrey, 2004). Gleason pattern 5 is assigned to cords and sheets of cells which have no discernible glandular structure and present comedonecrosis (glandular lumina filled entirely with necrotic cells). It represents the most invasive Gleason grade (Gleason and Mellinge.Gt, 1974). A descriptive summary, along with Dr. Gleason's original staging diagram, is provided in Figure 1.3.

The Gleason score, which is denoted as the combination or sum of the Gleason grades for the two largest areas of tumour foci in a specimen by order of size, technically ranges from 2 to 10, but rarely goes below 6 (Penney *et al.*, 2013), as histological findings are rarely assigned a Gleason pattern 2 (Melia *et al.*, 2006) and it is recommended that Gleason pattern 1 and 2 are never assigned to a prostate needle

biopsy (Gordetsky and Epstein, 2016). In addition, there are concerns regarding the poor inter-observer reproducibility of assigning a Gleason score of 2-4 (Allsbrook *et al.*, 2001), Gleason patterns 1, 2 and 3 are frequently mis-graded (Allsbrook *et al.*, 2001; Gordetsky and Epstein, 2016) and thus Gleason scores of 5-6 are often under-graded (Allsbrook *et al.*, 2001).

Another system used to categorise patients is tumour staging, which reflects the extent of tumour progression in the prostate and the surrounding organs, as well as factoring in lymph node and distant metastases. The current system in use is the three-parameter Tumour, Nodes and Metastases (TNM) staging system, which is used to guide patient treatment options (Cheng *et al.*, 2012). The T-stage denotes the extent of tumour involvement: T1 tumours are found incidentally upon transurethral resection of the prostate or are identified by needle biopsy but not palpable or visible by imaging techniques; T2 tumours are confined to the prostate but palpable or visible by imaging; T3 tumours have invaded locally and display extraprostatic extension, bladder neck invasion or seminal vesicle invasion; T4 tumours have more extensive local invasion into the rectum, pelvic wall or levator muscles (Cheng *et al.*, 2012). The N-stage is used to categorise cancer for pelvic lymph node metastases, with N1, N2 and N3 cancers having lymph node metastases of varying sizes and frequencies. Extrapelvic lymph node metastases are considered as M1, which is the category assigned to distant metastases to bone and other sites.

Evidence is unclear as to whether Gleason grade can change over time with increasing tumour stage. One study found that Gleason grade was not uniformly correlated with tumour stage (Penney *et al.*, 2013). However, an increase in Gleason grading or staging can occur between biopsy and surgery, as shown by a Japanese cohort of patients treated with prostatectomy and adjuvant hormone therapy (Mukouyama *et al.*, 2001). Gleason grading can also produce a higher grade for lymph node metastases than for the primary tumour in approximately 45% of cases, suggestive of dedifferentiation in a metastatic niche (Cheng *et al.*, 2012).

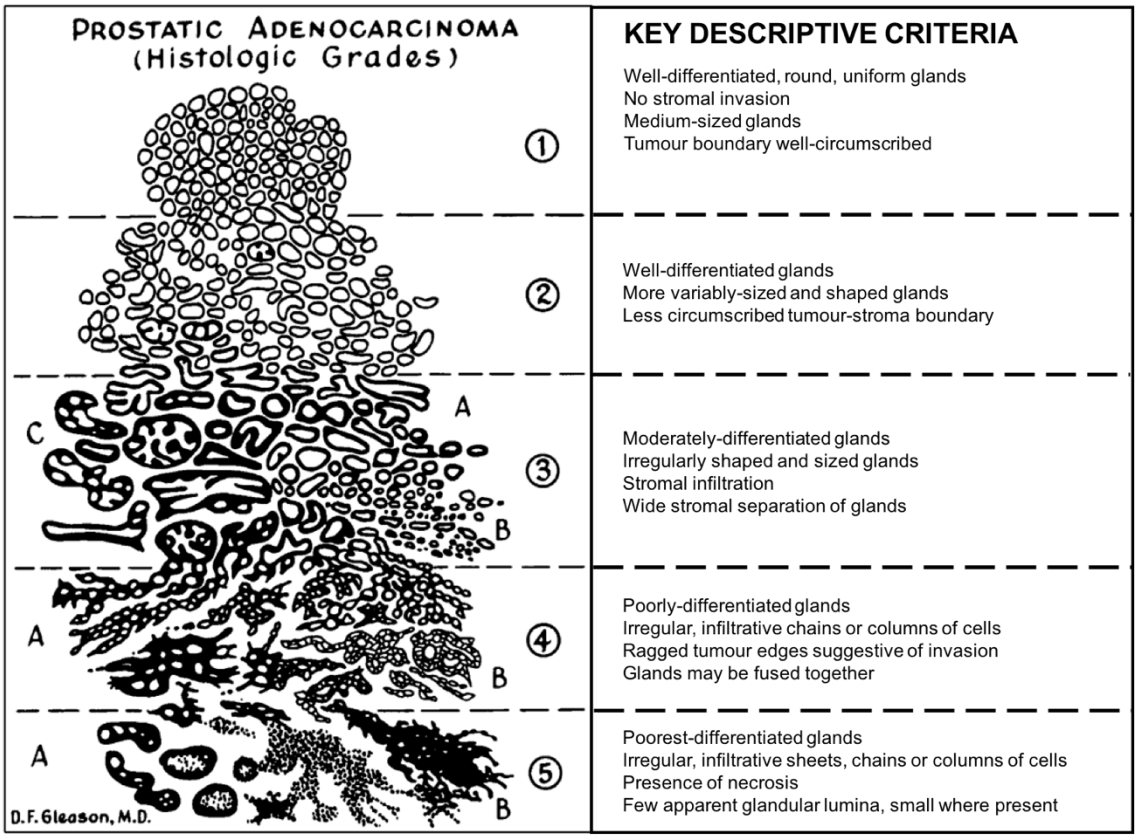


Figure 1.3: A graphical depiction of the Gleason scoring system drawn by Gleason, along with some key descriptive criteria.

Gleason grades are assigned with increasing degrees of glandular deformity from 1 to 5, with 1 being well-differentiated and 5 being poorly differentiated. Some of the key criteria for grading include: size, shape and uniformity of glands; presence or absence of glandular lumen; infiltration into the surrounding stroma and presence or absence of necrosis. This is not an exhaustive list and there are other factors for pathologists to take into account when Gleason scoring. Figure is adapted from Humphrey, 2004 and was originally presented in Gleason and Mellinger, 1974.

Despite the caveats of Gleason scoring, it remains one of the most robust diagnostic and prognostic variables in clinical use for prostate cancer (Gordetsky and Epstein, 2016). A high Gleason score is a positive predictor of metastasis following prostatectomy (Antonarakis *et al.*, 2012). There have been several attempts to improve the Gleason system, both in terms of its quantification and its reliability. In intermediate Gleason score 7 (3+4 or 4+3) disease, recent data suggests that quantification of the amount of Gleason pattern 4 in a biopsy can be a useful prognostic tool (Deng *et al.*, 2016; Cole *et al.*, 2016) with patients bearing a higher burden of Gleason pattern 4 at greater risk of recurrence. Therefore, Gleason scores of 3+4 and 4+3 are not equal in terms of predicted outcome (Gordetsky and Epstein, 2016).

A new grading system has very recently been adopted which combines information from the Gleason grading system into categories that better reflect the risk of recurrence for patients with prostate carcinoma (Epstein *et al.*, 2016a; Epstein *et al.*, 2016b). The new grading system creates 5 new Grade Groups: Grade Group 1 consists of any tumour with a Gleason grade of 6 or less; Grade Group 2 represents Gleason grade 3+4 tumours; Grade Group 3 represents Gleason grade 4+3 tumours; Grade Group 4 represents Gleason Grade 4+4 tumours; Grade Group 5 represents Gleason grades 9 and 10 (Epstein *et al.*, 2016b). The 5-year biochemical recurrence-free progression probabilities for the Grade Groups were 96%, 88%, 63%, 48% and 26% respectively, highlighting both the extraordinary clinical and morphological heterogeneity of prostate cancer, as well as the discriminatory power of this tool for prognostic use. However, there are only a limited number of other tools for diagnosis and prognostication in clinical use. Advances in immunohistochemistry, which was not originally part of the routine diagnosis of adenocarcinoma (Gordetsky and Epstein, 2016), form a large part of this technology and are discussed in the next section.

1.2.2.3 The role of IHC in prostate pathology

Histology is the workhorse of diagnostic pathology which aims to study the morphology of tissues and tissue components, using stains to bring out the contrast on molecules or cells of interest in the tissue to be viewed under a microscope. Recent advances in tissue fixation and specimen processing mean that special stain histology, IHC, fluorescent in-situ hybridisation and other nucleic acid and protein-based techniques can be used in tandem to provide good quality data at multiple levels: DNA, RNA, protein and histological features may all be evaluated using clinical specimens (Braun *et al.*, 2011).

These advances mean that there is the potential for incorporation of routine histological practice into an integrative approach, making use of the increasingly sophisticated molecular profiling techniques available at the transcriptomic and proteomic levels as these technologies become cheaper.

The current use of histology in prostate pathology concerns elements of both diagnosis and prognosis. The chief role of H&E staining is to morphologically characterise core needle prostate biopsies and transurethral resections of the prostate (TURP) to determine the extent and Gleason score of any tumours found (Varma *et al.*, 2002; Epstein, 2004), as well as to check surgical margins of an excised tumour for invasive cells. Diagnosis is often difficult with the small amount of material provided in a biopsy, as the wider glandular context is not evident and pathologists often have only a small focus of tumour on which to form a diagnosis (Varma *et al.*, 2002; Paner *et al.*, 2008; Epstein, 2004). There are many benign mimics of prostate cancer, including adenosis and incidental sampling of the nearby Cowper's glands or seminal vesicles (DeMarzo *et al.*, 2003), making a differential diagnosis difficult . A key feature of prostate carcinoma is the absence of basal cells, a discovery first made in 1953 (Totten *et al.*). Basal cells are not always easy to identify on an H&E stained tissue section, as the tangential cutting of other cells, such as luminal cells or fibroblasts, can mimic basal cell morphology (Varma *et al.*, 2002). Therefore, a differential diagnosis can be assisted by the use of IHC to identify these basal cells, or the absence thereof.

1.2.2.4 Basal cell markers

As mentioned in Chapter 1.1.2, basal cells express high molecular weight cytokeratins (HMWCK) such as CK5 and 14, which are routinely identified in IHC using an antibody against high molecular weight cytokeratins called 34BE12 (De Marzo *et al.*, 1999; Varma and Jasani, 2005). Additionally, basal cells express the p53 homologue p63 in the nuclear compartment (Signoretti *et al.*, 2000; Paner *et al.*, 2008). If desired, a stronger signal can be achieved by using an antibody cocktail with a mixture of antibodies identifying both p63 and HMWCK simultaneously (Paner *et al.*, 2008). The presence or absence of basal cells in a prostate biopsy can confirm a suspect lesion, as absence of basal cells is a defining characteristic of prostate carcinoma (Paner *et al.*, 2008; Varma and Jasani, 2005). Potentially pre-malignant lesions, known as high-grade Prostatic Intraepithelial Neoplasia (PIN), are sometimes difficult to distinguish from prostate carcinoma with an H&E stain (Epstein, 2004). By staining for basal cell

markers, it is possible to distinguish PIN from carcinoma, as PIN lesions retain a basal cell layer despite a crowded glandular epithelium (Torabi-Nezhad *et al.*, 2016), although the basal layer becomes increasingly discontinuous and disrupted with higher grades of PIN (Bostwick and Brawer, 1987; Liu *et al.*, 2009; Brimo and Epstein, 2012). Given the loss of basal cell continuity in high-grade PIN foci, and the small amount of tissue being assessed, there is always the possibility of a false positive diagnosis of carcinoma from this information alone (Varma and Jasani, 2005). This means that additional sources of information are required for diagnosis.

1.2.2.5 Tumour cell markers

Markers of tumour cells may assist in the differential diagnosis of carcinoma from PIN lesions. The most promising prostate tumour cell marker in current pathology use is α-methylacyl coenzyme A racemase (AMACR), or P504S. AMACR is an enzyme involved in the beta-oxidation of branched chain fatty acids, and is overexpressed at the mRNA (Luo *et al.*, 2002; Petrovics *et al.*, 2005; Ashida *et al.*, 2004; Grasso *et al.*, 2012) and protein level in prostate tumours compared to high grade PIN and benign prostate epithelium (Luo *et al.*, 2002). AMACR is either negative or weakly positive in the cytoplasmic compartment of benign prostate epithelial cells, is expressed at an intermediate intensity variably in PIN, and stains intensely in most prostate carcinoma specimens (Luo *et al.*, 2002). A 22-study meta-analysis of AMACR expression has shown that AMACR expression by PCR or by IHC is associated with increased risk of prostate cancer diagnosis (Jiang *et al.*, 2013). AMACR is most useful in the identification of small lesions suspicious for carcinoma (Brimo and Epstein, 2012).

The fact that AMACR is expressed in both PIN and carcinoma means that it is most commonly used as part of an antibody cocktail, sometimes known as the triple or PIN cocktail, which includes antibodies against AMACR, p63 and 34BE12 in a double-chromogen system (Paner *et al.*, 2008; Varma and Jasani, 2005; Tolonen *et al.*, 2011). The use of a PIN cocktail increases the successful detection of carcinoma in small lesions as well as reducing pathologist workload (Tolonen *et al.*, 2011). However, AMACR staining should only be considered in the context of concurrent basal cell marker staining, and the expected pattern for a carcinoma would be positive for AMACR and negative for basal cell markers in the focus, whereas PIN might be positive for both (Paner *et al.*, 2008; Varma and Jasani, 2005). According to a study by Zhou *et al.* in 2004, a positive AMACR result on an atypical, basal cell-negative lesion

in a biopsy is likely to result in upgrading to a diagnosis of cancer in up to half of all cases (Zhou *et al.*, 2004). However, the group also found 18% of cancer cases that did not stain with an anti-AMACR antibody, suggesting that there is a potential for false negative diagnoses (Zhou *et al.*, 2004; Brimo and Epstein, 2012). This reinforces the conclusion that while AMACR possesses considerable diagnostic utility, there is no single marker specific to prostate carcinoma in prostate pathology.

1.2.3 Progression and key molecular drivers of prostate cancer

Despite prostate cancer being a disease of men predominantly over the age of 50 (Siegel *et al.*, 2016), incidental findings of prostate cancer and high grade PIN, a precursor to carcinoma, on autopsy have been identified in men as young as 30 (Soos *et al.*, 2005). This suggests that the natural history of prostate cancer and the initiation of pre-neoplastic events begin early on in life. There are two potential precursor lesions to carcinoma: Proliferative Inflammatory Atrophy (PIA) and Prostatic Intraepithelial Neoplasia (PIN).

PIA is a focal change in prostate architecture marked by increased presence of proliferation markers within prostatic epithelium (Feneley *et al.*, 1996), and an increased burden of inflammatory cells in stromal and epithelial prostate compartments (De Marzo *et al.*, 1999; De Marzo *et al.*, 2003). It is accompanied by cytological alterations in luminal cells, namely a reduced, basophilic- to clear-staining cytoplasmic compartment. Basal cell layers remain intact. Meanwhile, PIN features benign glands with remaining basal cell architecture but with nuclear atypia and crowding (Epstein, 2009). It is thought that the inflammatory nature of PIA drives progression to PIN, and PIN subsequently progresses to carcinoma, although this is controversial (De Marzo *et al.*, 2003; Abate-Shen and Shen, 2000).

The transition from benign prostate to PIN to carcinoma is marked by a number of key changes that support this hypothesis. Firstly, there is a progressive disruption and disappearance of the basal cell layer along with disruptions to the basement membrane underlying these cells: partial basal cell and basement membrane disruption is already apparent in PIN and is complete in carcinoma (Liu *et al.*, 2009). This disruption of the basal cell layer is also apparent in several transgenic mouse models of PIN where the tumour suppressor PTEN is deleted, a constitutively active myristoylated form of AKT is expressed, or the oncogene c-myc is expressed (Majumder *et al.*, 2008). Secondly, there is a progressive increase in copy number alterations in PIA, PIN and carcinoma,

an example being chromosome 8 gains in 8q24, which contains the oncogene c-myc (Yildiz-Sezer *et al.*, 2006) and is often progressively overexpressed during these transitions (Tomlins *et al.*, 2007). In addition, PIN and prostate carcinoma share similar patterns of global gene expression, as identified by shared expression signatures (Tomlins *et al.*, 2007). A series of defined genetic alterations and signalling aberrations appear to underpin the transitions from PIA to PIN and then to carcinoma, which will be discussed in this section. Each molecular alteration or signalling pathway will be discussed individually, although where possible the order in which these processes contribute to initiation and progression is preserved. A summary of key driver events in prostate carcinogenesis is presented in Figure 1.4.

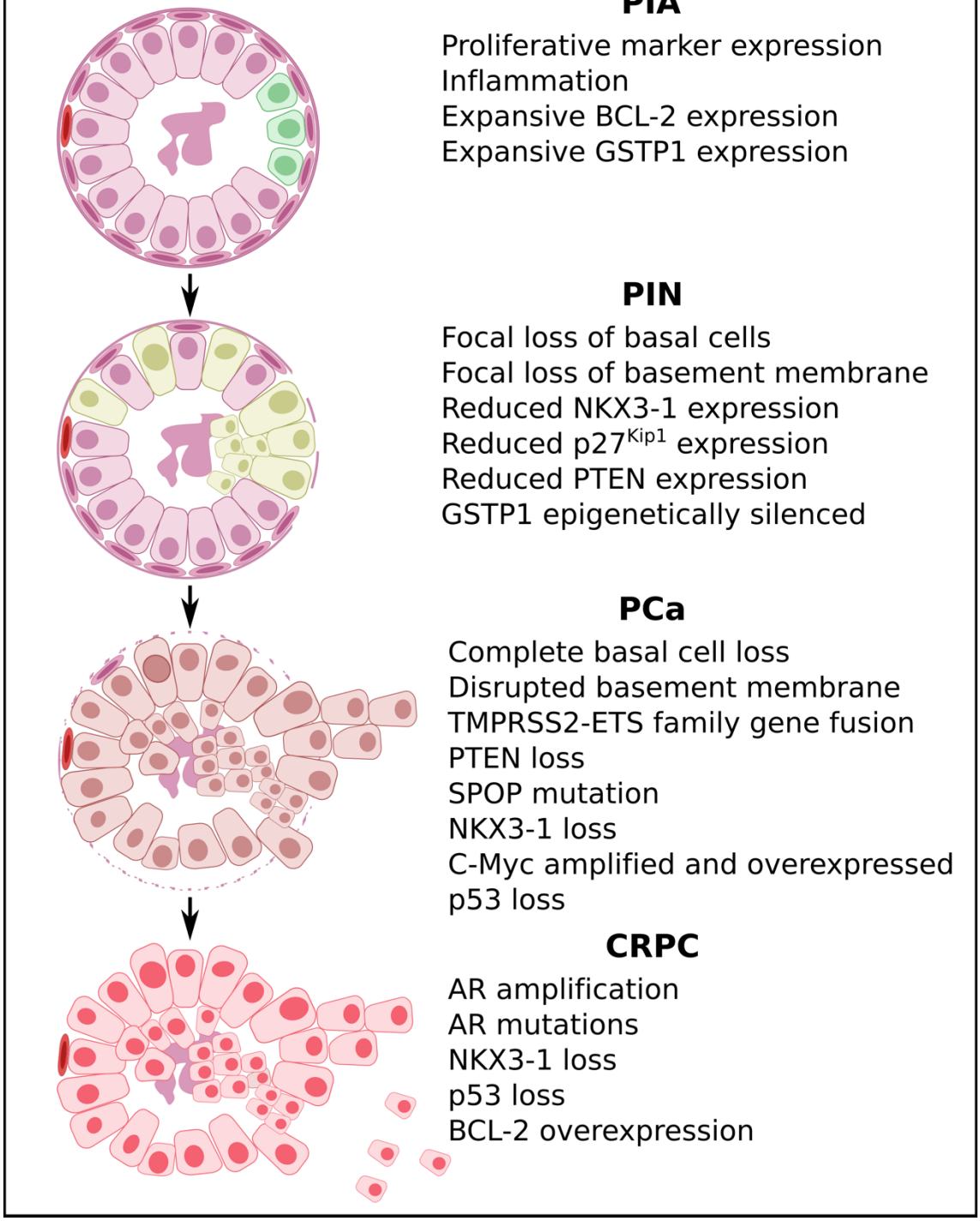

PIA
Proliferative marker expression
Inflammation
Expansive BCL-2 expression
Expansive GSTP1 expression
PIN
Focal loss of basal cells
Focal loss of basement membrane
Reduced NKX3-1 expression
Reduced p27Kip1 expression
Reduced PTEN expression
GSTP1 epigenetically silenced
PCa
Complete basal cell loss
Disrupted basement membrane
TMPRSS2-ETS family gene fusion
PTEN loss
SPOP mutation
NKX3-1 loss
C-Myc amplified and overexpressed
p53 loss
CRPC
AR amplification
AR mutations
NKX3-1 loss
p53 loss
BCL-2 overexpression

Figure 1.4: A summary of known key oncogenic events in prostate carcinogenesis.

Key molecular events, including gene mutations, losses and rearrangements, are presented next to a visual representation of each stage of prostate carcinogenesis. The stages of prostate carcinogenesis in this model are: Proliferative Inflammatory Atrophy (PIA, green cells); Prostatic Intraepithelial Neoplasia (PIN, yellow cells); Prostate Carcinoma (PCa, brown cells); and Castration-Resistant Prostate Carcinoma (CRPC, red cells).

1.2.3.1 GSTP1 expression is lost from PIN onwards

The oxidative stress response protein glutathione S-transferase π (GSTP1), which is normally expressed in basal cells in normal prostate glands, is expressed more expansively in PIA, focally including luminal epithelial cells in addition to its normal basal staining pattern (De Marzo *et al.*, 1999). This is accompanied by increased expression of the anti-apoptotic protein Bcl-2. Interestingly, CpG methylation at the GSTP1 promoter is also a frequent occurrence during tumorigenesis and occurs as early as PIN (Kang *et al.*, 2004) and through to carcinoma (Kang *et al.*, 2004; Abeshouse *et al.*, 2015), resulting in downregulation of GSTP1 protein levels (Lin *et al.*, 2001). Perhaps oxidative stress selects for new mutations in pre-neoplastic cells to facilitate progression. It would suggest that oxidative stress, apoptosis, inflammation and the appearance of potentially pre-neoplastic foci in the prostate may be linked together. However, the status of PIA as the precursor lesion to PIN is still controversial, since there are very few common genetic mutations between the two on a global level (Tomlins *et al.*, 2007).

1.2.3.2 NKX3-1 is downregulated in PIN and deleted in carcinoma

Homeobox protein NKX3-1 is a transcription factor first expressed in the developing endodermal compartment known as the urogenital sinus, expressed at 15.5dpc before the prostate begins to form, and is therefore one of the first markers of the developing prostate (Abate-Shen *et al.*, 2008). Its expression is dependent on androgen signalling, and it is required for secretory differentiation and branching morphology in the mouse prostate. Chromosomal alterations at 8p12-21 are thought to occur early in the process of prostate carcinogenesis (Abate-Shen and Shen, 2000) and include loss of the luminally-expressed tumour suppressor NKX3-1 at 8p21.2. Deletion of even a single NKX3-1 allele in mouse models generates PIN lesions which lack basal cells (Abdulkadir *et al.*, 2002), and transgenic adenocarcinoma of the mouse prostate

(TRAMP) mice display progressive loss of NKX3-1 expression as they progress to invasive carcinoma (Bethel and Bieberich, 2007). NKX3-1 deficient lesions have progressive oxidative damage to DNA and protein along with concurrent decreases in expression of several antioxidant enzymes, suggesting that a key function of the protein in mice is to regulate the response to oxidative stress, which could have a role in the initiation of cancer (Ouyang *et al.*, 2005).

In the human prostate, modest reductions of NKX3-1 expression are observed in early pre-neoplastic events such as PIA, and the downregulation carries through to PIN and carcinoma (Bethel *et al.*, 2006). However, losses of heterozygosity at the 8p locus are only found in a 12% of high-grade PIN lesions, while the deletion frequency increases to 33% and 74% of Gleason pattern 3 and Gleason pattern 4/5 tumours respectively (Bethel *et al.*, 2006). In a separate study, 20% of high-grade PIN lesions were identified as having completely lost NKX3-1 expression, while in stage T1, T2 and T3/4 tumours this was 6%, 16% and 22% respectively (Bowen *et al.*, 2000). Loss of NKX3-1 expression was most prominent in castration-resistant and metastatic tumours with a frequency of 34% and 78% respectively. Mutations in NKX3-1, however, are rare (Abeshouse *et al.*, 2015). Taken together, the evidence suggests that downregulation of NKX3-1 expression is an early event in prostate carcinogenesis, as it is detected as early as PIA and is observed through to carcinoma, although the mechanism for downregulation is still not understood (Abate-Shen *et al.*, 2008). However, the more complete losses of NKX3-1 expression and losses of heterozygosity at the locus are more likely to be associated with progression to advanced or metastatic disease (Bowen *et al.*, 2000).

1.2.3.3 $p27^{Kip1}$ is downregulated in the PIN-to-carcinoma transition

As mentioned at the beginning, there are several mouse models for PIN: prostate-specific PTEN deletion, myristoylated AKT expression, or c-myc expression will all generate PIN (Majumder *et al.*, 2008). Oncogenic signalling in these models results in increased expression of the cyclin-dependent kinase inhibitor $p27^{Kip1}$, which induces senescence and acts as a brake to proliferation in the epithelial compartment of PIN lesions. Thus, progression to carcinoma does not happen in these models unless $p27^{Kip1}$ expression is perturbed (Majumder *et al.*, 2008; Wang *et al.*, 2012a; Di Cristofano *et al.*, 2001) and this is thought to be due to changes in cell adhesion in the PIN environment

rather than from aberrant PI3K-Akt signalling induced by PTEN (Majumder *et al.*, 2008). This highlights the importance of tumour-stromal interactions in prostate cancer.

In the context of human prostate tissue, $p27^{Kip1}$ levels are suppressed in PIN lesions adjacent to tumours, whereas they are maintained in PIN lesions not associated with cancer (Majumder *et al.*, 2008). Similarly, a reduction in staining for the cell cycle regulator $p27^{Kip1}$ was observed in prostate tumour tissues when compared to benign controls and PIN in a separate cohort (Tsihlias *et al.*, 1998). In contrast, other studies have found that $p27^{Kip1}$ was weakly expressed in almost all studied PIN lesions (Fenic *et al.*, 2004), and that $p27^{Kip1}$ expression was just as low in PIN as in carcinoma (Doganavsargil *et al.*, 2006). However, reduced expression of $p27^{Kip1}$ correlates with recurrence in patients with primary prostate carcinoma and especially in those receiving neoadjuvant hormone therapy (Tsihlias *et al.*, 1998), which could suggest an association between loss of $p27^{Kip1}$ expression and castration resistance. While the resolving power of $p27^{Kip1}$ expression for PIN and carcinoma is unclear, it is clear that in mice and in humans, the progression from PIN to invasive carcinoma requires the deregulation of cell cycle control and senescence programs, such as those regulated by $p27^{Kip1}$ and p53 (see Chapter 1.2.3.6), which appear to be initiated in the context of PIN and progression-induced changes of prostate epithelial organisation.

1.2.3.4 C-myc amplification and overexpression occur throughout prostate carcinogenesis

C-myc is a transcription factor oncogene located on chromosome 8q24.21 that is frequently amplified and overexpressed in many human tumours, including prostate cancer (Dang, 2012). It lies downstream of WNT signalling and many receptor tyrosine kinases, which positively regulate it at the transcriptional level (Dang, 2012). In combination with its co-activator max, c-myc activates gene expression and promotes cell growth and proliferation, while also being one of the Yamanaka factors necessary for reprogramming of differentiated cells to an embryonic stem cell-like state (Dang, 2012). C-myc can also be regulated by microRNAs: miRNA-34a, a negative regulator of cancer stem cell phenotypes (see Chapter 1.3.6) and a transcriptional target of p53, directly represses c-myc at the translational level during oncogene-induced senescence (Christoffersen *et al.*, 2010).

Mechanistically, c-myc expression in prostate cancer is important in the control of cell growth and proliferation. A microarray study of myc overexpression in fibroblasts

revealed that c-myc both represses and promotes transcription of a diverse set of target genes: it represses cell cycle inhibitor p21 and promotes expression of Cyclin D2, suggesting that it promotes cell cycle progression (Coller *et al.*, 2000) as well as the increased expression of translation initiation factors eIF4 γ and eIF5E, to name a few. Consistent with these roles, in the prostate cancer cell line LNCaP, c-myc overexpression upregulates biosynthetic pathways necessary for cell growth including ribosomal biogenesis and purine biosynthesis, and in castration-resistant prostate tumours it tightly correlates with expression of an enzyme involved in guanine nucleotide biosynthesis, IMPDH2 (Barfeld *et al.*, 2015). Inhibition of IMPDH2 activity or downregulation with siRNA in LNCaP cells resulted in reduced cell proliferation which synergised with anti-androgens (Barfeld *et al.*, 2015), suggesting that targeting the c-myc-driven regulation of biosynthesis could be therapeutically effective in castration-resistant prostate carcinoma.

Mice that overexpress c-myc in a prostate-specific manner develop PIN followed by progression to carcinoma between 6 and 12 months of age, and are one of the few mouse models of prostate cancer that do not require ablation of a tumour suppressor gene to progress to carcinoma (Ellwood-Yen *et al.*, 2003). Expression of NKX3-1 at the protein level is reduced in PIN and then lost in carcinoma, reminiscent of events in human prostate carcinogenesis (Ellwood-Yen *et al.*, 2003). However, as with most mouse models of prostate cancer, they fail to be fully representative of the human disease: there were no genomic alterations underlying the transformation to PIN or the progression to carcinoma, which are pivotal in progression of the human disease (Ellwood-Yen *et al.*, 2003).

Evidence for a role of c-myc in human prostate cancer initiation and progression is strongest at the DNA level. In a FISH study, amplification of c-myc is observed in approximately 10% of metastatic tumours and in 5% of locally recurrent tumours (Bubendorf *et al.*, 1999). Another FISH study found gains in 8q24 in 19% of PIA lesions, 21% of PIN lesions and in 27% of carcinoma lesions (Yildiz-Sezer *et al.*, 2006). Interestingly, in this study even 10% of normal prostate tissue had amplification of c-myc: this study either suggests that 8q24 amplifications occur very early on in prostate tumorigenesis, or there could be sources of false-positive amplification events being counted, such as overlapping nuclei in the sections. However, copy number gains of the whole of chromosome 8, which contains the c-myc locus, are strikingly common during prostate tumorigenesis (Jenkins *et al.*, 1997; Sato *et al.*, 2006): when factoring in these

events, c-myc amplifications occur in approximately 50% of PIN lesions, 44% of carcinoma lesions and 92% of metastases (Jenkins *et al.*, 1997). In locally advanced stage (T3) prostate tumours treated with radical prostatectomy, copy number gain of c-myc is associated with a higher Gleason score, and the presence of additional copy number gains relative to the copies of chromosome 8 is significantly associated with an earlier disease recurrence (Sato *et al.*, 2006).

Furthermore, c-myc mRNA expression in primary prostate tumours seems to be prognostically important, as it is an independent predictor of biochemical recurrence (Hawksworth *et al.*, 2010). Using a qRT-PCR-based assay, Hawksworth et al. found that c-myc expression was positively correlated with expression of ETS family oncogene ERG (Hawksworth *et al.*, 2010) (see Chapter 1.2.3.6). In addition, concurrent ERG and c-myc expression at the protein are associated with locally advanced prostate carcinoma (Udager *et al.*, 2016), potentially indicating a tumour subtype-specific effect of c-myc on prostate cancer behaviour.

Interestingly, c-myc mRNA levels were not observed to correlate with c-myc protein levels (Hawksworth *et al.*, 2010), perhaps suggesting that c-myc is subject to additional regulation at the mRNA level as detailed previously. At the protein level, normal, PIA, PIN and carcinoma specimens are positive for c-myc in 23%, 21%, 76% and 82% of cases respectively (Gurel *et al.*, 2008). Combined with evidence at the copy number level, this suggests that c-myc exerts its diverse effects on carcinogenesis throughout the spectrum from PIN to carcinoma and metastases. Although copy number changes and mouse models suggest that it could be a driver event for PIN-to-carcinoma transition, it is also implicated in tumour recurrence and heavily amplified in castration-resistant disease. C-myc-mediated signalling likely collaborates with other driver mutations in the maintenance of a castration-resistant phenotype, promoting cell cycle progression and cell growth.

1.2.3.5 Other alterations during PIN-to-Carcinoma progression

Other commonalities between PIN and prostate cancer include upregulation of AMACR and spermine synthase expression, and downregulation of complement components and serine protease inhibitors of the SERPIN family (Ashida *et al.*, 2004). Synthesis of polyamines such as spermine are consistently elevated in prostate carcinoma (Rhodes *et al.*, 2002), and their depletion results in a growth inhibitory effect in mouse prostate adenocarcinoma models and *in vitro* (Kee *et al.*, 2004). Similarly, depletion of

complement components is observed in prostate carcinoma and BPH specimens compared to normal tissue (Hong *et al.*, 2009), and complement factor C1q promotes apoptosis in prostate cancer cell lines in vitro, a process mediated by the tumour suppressors p53 and WOX1. This suggests that a key process in the transformation from PIN to carcinoma is the suppression of anti-proliferative and pro-apoptotic signals.

With respect to SERPINs, a SERPIN family member known as maspin is frequently downregulated in prostate cancer, and has been associated with shorter recurrence-free survival in prostate cancer patients (Machtens *et al.*, 2001). Additionally, expression of maspin in PC-3 and DU145 prostate cancer cell lines sensitises them to treatment with the chemotherapeutic gemcitabine (Huang *et al.*, 2016), and in DU145 xenografts it also results in recruitment of B-cells and neutrophils accompanying an anti-tumour immune response (Dzinic *et al.*, 2014). Interestingly, a cluster of SERPINs on 18q21 are frequently deleted in prostate cancer (Machtens *et al.*, 2001; Yildiz-Sezer *et al.*, 2006), which might explain the reduced expression of this protein family. This reflects an imbalance in proteolysis in favour of a more proteolytic prostatic environment (Machtens *et al.*, 2001) and is another of the myriad copy number changes which are common on the route from precursors lesions to PIN and prostate carcinoma, albeit with progressively increasing frequency (Yildiz-Sezer *et al.*, 2006).

1.2.3.6 ETS gene family members fuse with TMPRSS2 in prostate carcinoma

Gene expression signatures identify key changes in gene expression in the transition from PIN to carcinoma, and these include increased expression of E26 transformation-specific (ETS) transcription factor family target genes, which is likely linked to increased protein biosynthesis (Tomlins *et al.*, 2007). TMPRSS2-ERG gene fusion is the most common ETS family alteration in prostate cancer, where the ETS family oncogene ERG becomes expressed in an androgen-responsive manner due to fusion with the TMPRSS2 promoter. It is a common gene fusion event in prostate cancer, occurring in approximately 50% of cases (Wyatt *et al.*, 2014; Abeshouse *et al.*, 2015; Taylor *et al.*, 2010). The TMPRSS2 and ERG loci, which are normally 3MB apart on chromosome 21q22.2, are brought into closer physical proximity with one another in androgen-dependent prostate cancer cells by ligand-bound AR (Mani *et al.*, 2009), and TMPRSS2-ERG gene fusion events similar to those observed in human tumours can be brought about by the introduction of DNA double strand breaks. This suggests that a mutational environment, in the context of active androgen receptor signalling,

contributes to the extraordinarily high level of ETS translocation events in prostate cancer.

Other ETS family members, ETV1 (Tomlins *et al.*, 2005; Abeshouse *et al.*, 2015), ETV4 and FLI1 (Abeshouse *et al.*, 2015), are also subject to fusions with TMPRSS2, though less frequently. Collectively, overexpression of ETS family genes are found in 50% to 70% of all primary tumours (Tomlins *et al.*, 2007; Tomlins *et al.*, 2005; Abeshouse *et al.*, 2015) and are specific to prostate carcinoma, not being present in PIN lesions (Tomlins *et al.*, 2005). Together they serve as useful biomarkers for molecular classification of prostate tumours: a recent study identified that 74% of primary prostate tumours could be separated into seven distinct molecular subtypes, four of which were based on ETS fusion events and comprised 59% of all tumours (Abeshouse *et al.*, 2015).

Interestingly, despite high alteration frequency, the presence of ERG gene fusion is a predictor for poor outcome in patients with prostate carcinoma (Sboner *et al.*, 2010). However, monoallelic TMPRSS2-ERG fusion has been associated with lower Gleason score tumours, whereas concurrent TMPRSS2-ERG and SLC45A3-ERG fusion was likely to increase ERG expression, co-occur with PTEN loss, and was associated with higher Gleason grade and stage tumours (Hernandez *et al.*, 2016). This suggests that ERG gene fusion precedes other key molecular changes which lead to an aggressive phenotype, although further studies of SLC45A3-ERG fusion would be needed to confirm this.

In mice, overexpression of transgenic ERG under the control of androgen responsive promoters is sufficient to induce PIN, but it is not sufficient for progressing to adenocarcinoma (Tomlins *et al.*, 2008; Lawson *et al.*, 2010; Chen *et al.*, 2013). As with many molecular alterations in the mouse prostate, it is hypothesised that more oncogenic insults are necessary to drive oncogenic transformation. It is important to remember that ERG fusion events occur in the context of complex oncogenic stimuli during tumour development, rather than in isolation. ERG overexpression appears to have no effect on proliferation, but does drive an invasive programme even in benign prostate cell line RWPE and primary prostate epithelial cells, thought to be through plasminogen activator-dependent activation of matrix metalloproteinases (Tomlins *et al.*, 2008). ERG expression in radical prostatectomies and metastases has also been associated with NF-kB activation and modulation of immune cell migration at the

transcriptome level by microarrays, and confirmed at the protein level by IHC (Roudier *et al.*, 2016). ERG expression is correlated with the expression of a stem cell-related marker DLCK1 in both hormone-naïve and castration-resistant prostate carcinoma (Roudier *et al.*, 2016). ERG expression in transgenic mice appears to positively regulate binding of AR to its transcriptional targets through a pioneer effect rather than through direct co-recruitment, and increases AR target gene transcription (Chen *et al.*, 2013). Other ETS transcription factors such as ETV1 also have this effect (Chen *et al.*, 2013). This suggests that ERG plays a role in upregulating AR signalling during the initial androgen-dependent phase of prostate carcinoma. In ERG mice and in human prostate tumours and cancer cell lines, deletion of PTEN causes a progression to invasive carcinoma and compensates for the reduction in androgen-mediated transcription that occurs as a result of PTEN loss (see Chapter 1.2.3.8) (Chen *et al.*, 2013).

Taken together, ETS fusion events probably occur early in prostate carcinoma development, after the transition from PIN to carcinoma. Such commonly mutated genes are clearly likely to serve a purpose in carcinoma development, and although the biological consequences of ETS fusion events remain to be completely understood, they appear to be important in priming the prostate for further oncogenic alterations.

1.2.3.7 P53 mutations and deletions occur in the transition to advanced stage disease

p53 is a key integrator of multiple signals including DNA damage, hypoxia and excessive oncogenic signalling, oversees multiple cell cycle checkpoints and also functions in the induction of apoptosis and senescence (Levine, 1997). This makes it a key tumour suppressor of prostate carcinoma. In mice, PIN lesions of PTEN-null transgenic models may progress to invasive carcinoma after the inactivation of p53, which otherwise induces senescence (Chen *et al.*, 2005) suggesting that p53-mediated cell cycle regulation is important in the PIN-to carcinoma transition. This has yet to be proven in humans, where studies suggest that it is an important driver in the transition to advanced stage prostate carcinoma (Navone *et al.*, 1993; McDonnell *et al.*, 1997; Meyers *et al.*, 1998; Qian *et al.*, 2002; Ecke *et al.*, 2010; Kuczyk *et al.*, 1998; Machtens *et al.*, 2001).

p53 deletion occurs in approximately 30% of advanced-stage primary prostate tumours (Qian *et al.*, 2002; Ecke *et al.*, 2010), where it is associated with an increased risk of recurrence and a shorter time to recurrence (Ecke *et al.*, 2010). In high Gleason grade androgen-independent prostate tumours (Navone *et al.*, 1993) and in metastases

(McDonnell *et al.*, 1997; Meyers *et al.*, 1998), mutation of the locus results in nuclear accumulation of p53, and missense mutations of p53 are associated with increased p53 stability (Navone *et al.*, 1993). Such alterations are not found in benign prostate tissue or in PIN lesions, suggesting that the p53 alteration is an event occurring on transition to advanced disease and metastasis (Navone *et al.*, 1993). This indicates that perturbation or alteration of p53 function is important for tumour progression to castration-resistant prostate cancer (CRPC) or metastasis.

However, exon-sequencing and copy number alteration studies conducted by Taylor *et al.* suggest that missense mutation of p53 is a rare event, while deletion of the locus is much more common at 24% of tumours (Taylor *et al.*, 2010), which agrees with previous studies on p53 deletion (Qian *et al.*, 2002). Interestingly, while p53 is a tumour suppressor and is frequently mutated and deleted in prostate cancer, it is also overexpressed in some primary tumours (Kuczyk *et al.*, 1998), and p53 overexpression is associated with increased risk of recurrence following prostatectomy (Kuczyk *et al.*, 1998; Machtens *et al.*, 2001). While p53 clearly has roles to play in prostate cancer progression to advanced stages, it appears that they are context-dependent, and currently incompletely understood.

1.2.3.8 Aberrations in PI3K-Akt signalling occur on the route to castration resistance

The phosphatidylinositol 3 kinase (PI3K) pathway is a signalling pathway that begins at the cell membrane and serves to transduce RAS and growth factor receptor signalling (Keniry and Parsons, 2008). Activated RAS or growth factor receptors in turn activate PI3K, which acts to phosphorylate the lipid phosphatidylinositol bisphosphate (PIP2) to produce PIP3. PIP3 is recognised by some proteins containing the pleckstrin homology domain, and acts to recruit and activate these proteins at the plasma membrane. The most notable of these is the kinase AKT, which mediates proliferation and cell survival (Keniry and Parsons, 2008). Phosphatase and tensin homologue (PTEN) acts to inhibit this signalling pathway by dephosphorylating PIP3 to PIP2 and directly antagonising PI3K activity (Stambolic *et al.*, 1998).

Mice with mutant PTEN have increased cell proliferation and reduced sensitivity to apoptosis (Stambolic *et al.*, 1998). PTEN deletion predisposes to PIN formation in mice (Majumder *et al.*, 2008; Chen *et al.*, 2006), and deletion of AKT1 is sufficient to inhibit this process, suggesting that the transformative event was brought about by aberrant AKT signalling (Chen *et al.*, 2006). The PTEN protein interaction network is also

significantly mutated in human prostate cancer (Grasso *et al.*, 2012; Abeshouse *et al.*, 2015). Alterations of PI3K signalling, either through over- and under-expression of signalling components or through somatic mutation, are found in approximately 42% of primary tumours and in 100% of metastases in a large-scale genomic and transcriptomic characterisation of prostate tumours (Taylor *et al.*, 2010). The transition from PIN to carcinoma appears to be marked by loss of expression of PTEN (Torabi-Nezhad *et al.*, 2016; Lotan *et al.*, 2013). Primary tumours not only have PTEN deletions in approximately 17% of cases (Abeshouse *et al.*, 2015; Berger *et al.*, 2011; Ahearn *et al.*, 2016), but also have a low frequency of activating mutations in downstream signalling components such as PIK3CA and PIK3CB (subunits of PI3-Kinase) and AKT1 (Abeshouse *et al.*, 2015).

Interestingly, in mouse models which express a homozygous TMPRSS2-ERG fusion transgene, invasive carcinoma ensues only when homozygous deletions of PTEN are introduced (Chen *et al.*, 2013). As mentioned before, upregulation of ETS gene expression in prostate epithelium in the context of PTEN loss restores some of the lost AR signalling frequently observed in PTEN-negative prostate tumours (Chen *et al.*, 2013). PI3K-AKT and AR signalling exhibit bidirectional cross-talk in the context of PTEN-negative prostate tumours: AKT inhibition restores AR signalling, whereas AR inhibition restores PI3K signalling (Carver *et al.*, 2011). This is thought to be due to reciprocal negative feedback loops, where an androgen-dependent AKT phosphatase PHLPP is positively regulated by AR, and PI3K-AKT signalling negatively regulates levels of the receptor tyrosine kinases HER2/3, which stabilise AR and promote its transcriptional activity (Carver *et al.*, 2011). This would help to explain why PI3K-AKT signalling persists after anti-androgens or castration, and why combined PI3K and AR inhibition is a more effective therapy than either alone (Zhang *et al.*, 2009). Similarly, NKX3-1;PTEN mutant mice develop castration resistance early (Gao *et al.*, 2006), suggesting that PTEN deletion could be an escape route for these cells to become castration resistant. This is consistent with a recent finding, in which PTEN loss is associated with lethal progression in patients with primary prostate carcinoma in the context of ERG-negative tumours (Ahearn *et al.*, 2016), but not in ERG-positive tumours (where AR signalling is presumably sustained). Clearly, progression to an androgen-independent state is dependent on multiple factors, but the appearance of PI3K aberrations at both early and late stages of prostate tumour development underlines the importance of this signalling pathway in multiple stages of the disease.

1.2.3.9 AR signalling is key to castration resistance

Prostate epithelial cells are androgen-dependent and the resulting tumours are normally treated with hormone therapy once primary interventions have failed (Heidenreich *et al.*, 2008). When comparing gene expression signatures of PIN and carcinoma, increased androgen signalling is observed in carcinoma (Tomlins *et al.*, 2007), suggesting an increased reliance on AR-mediated signalling outputs in the primary disease. In the absence of androgens, tumour cells are under selective pressure to survive and proliferate, resulting in numerous alterations to signalling and either the generation of, or selection for, a castration-resistant population of cancer cells (Grasso *et al.*, 2012). Despite the low level of androgens in a castrate environment, most prostate tumours continue to rely on AR signalling (Groner *et al.*, 2016). AR signalling has a diverse set of functions, including in cell cycle progression and in steroid, fatty acid and carbohydrate metabolism (Massie *et al.*, 2011). Therefore, there are possibly many routes to castration resistance and metastasis, but AR signalling is nevertheless central to this process. This section will focus on the roles of AR signalling in the development of castration-resistant disease.

Development of castration resistance may occur early on in prostate cancer progression, with androgen-independent cancer cells being identified in early primary tumours through tissue culture approaches (Finones *et al.*, 2013) and it is hypothesised that this pre-existing 'tumour-initiating' cell/cancer stem cell (CSC) population could be selected for during androgen deprivation therapy to drive castration resistant disease (see Chapter 1.3 for more information on tumour-initiating cells). Consistent with this argument, in prostate cancer patients and in the prostate cancer cell line DuCaP the acquisition of castration resistance results in the upregulation of AR, steroidogenic enzymes and CSC markers such as CD44 (Pfeiffer *et al.*, 2011). BCL-2 expression is also elevated in patients with castration resistant tumours as compared to hormone-sensitive tumours, consistent with the concept that tumour cells are selected for their ability to resist androgen withdrawal-induced apoptosis (McDonnell *et al.*, 1992). Similarly, castration resistance can manifest early in murine tumour development if the tumour suppressors NKX3-1 and PTEN are both deleted (Gao *et al.*, 2006), suggesting that PI3K-AKT signalling and the background of tumour suppressor deletion is important in castration resistance, and is particularly important due to the cross-talk with AR signalling, as mentioned in Chapter 1.2.3.8.

1.2.3.9.1 AR overexpression and mutation

One common set of alterations that arises in response to hormone therapy is alteration to the AR locus itself. AR amplification is a frequent event in the development of castration-resistant prostate cancer (Grasso *et al.*, 2012; Ulz *et al.*, 2016), arising after hormone therapy (Visakorpi *et al.*, 1995), presumably in order to sustain a level of AR signalling that will support survival and proliferation. Indeed, AR expression may be important for cell proliferation in both androgen-dependent and some apparently androgen-independent cells *in vitro* (Zou *et al.*, 2006). AR mutations are also observed, many of which arise as a result of selective pressures brought about by specific anti-androgen treatments (Steinkamp *et al.*, 2009). Sequencing of AR cDNA in tumours reveals that there are many diverse mutations in AR throughout the coding sequence, including seven recurring unique mutations induced by bicalutamide and eight induced by flutamide (Steinkamp *et al.*, 2009). Mutations in AR have many different effects that contribute towards treatment resistance: AR-V716M has promiscuous binding and is able to bind adrenal androgens and progesterone at higher affinity than wild type AR (Culig *et al.*, 1993), whereas H875 and T878 activate AR signalling in the presence of flutamide, and AR-E255K has increased stability and nuclear localisation even in the absence of bound ligand (Steinkamp *et al.*, 2009). Similarly, LNCaP prostate cancer cells that have been selected for resistance to bicalutamide develop a mutation in AR-W741 that allows them to use bicalutamide as an agonist (Hara *et al.*, 2003), although AR is still inhibited by other anti-androgens such as hydroxyflutamide. This suggests resistance involves sustaining AR activity in the presence of AR inhibitors and castrate hormone levels, but that there are multiple means to achieve this. Indeed, deep sequencing of AR in castration-resistant prostate cancer circulating plasma DNA reveals that there are often multiple, heterogeneous tumour clones which are dynamically regulated in both a spatial and temporal manner in response to treatment (Carreira *et al.*, 2014). For example, treatment with different hormone therapies results in emergence and regression of clones with different AR mutations, such as a glucocorticoid-activated AR-H875Y mutant which arose in a patient in response to prednisolone treatment, regressed upon addition of the anti-androgen abiraterone, only to re-emerge with an additional AR-T878A mutation upon treatment failure (Carreira *et al.*, 2014). The dynamic acquisition and regression of different AR mutations in CRPC tumour clones due to selection explains why switching anti-androgens in a failing hormone therapy

regimen can sometimes prolong their effectiveness (Sternberg *et al.*, 2013; Heidenreich *et al.*, 2008).

1.2.3.9.2 Mutations in coactivators and regulators of AR signalling

Other routes to castration resistance involve sustaining AR signalling through mutations in other AR-interacting proteins. SPOP mutations, which result in increased AR target gene induction, have been recorded in primary prostate tumours (Abeshouse *et al.*, 2015). SPOP is a subunit of a cullin-type E3 ubiquitin ligase whose mutations are found in 10% to 13% of primary prostate tumours which are always ETS-fusion negative (Barbieri *et al.*, 2012; Abeshouse *et al.*, 2015) and appear to be located in the substrate-binding pocket of the enzyme, suggesting altered substrate binding affinity or specificity. Prostate cancer cells expressing mutant SPOP or SPOP siRNA had increased invasive capabilities *in vitro* (Barbieri *et al.,* 2012)*,* similar to findings with ERG expression. Interestingly, SPOP mutation is mutually exclusive with ERG fusion, suggesting two distinct molecular subtypes of cancer with different routes to castration resistance (Abeshouse *et al.*, 2015). SPOP has been shown to target SRC-3, a steroid receptor co-activator overexpressed in many cancer types, for proteasomal degradation (Li *et al.*, 2011). SRC-3 is a co-activator for AR (Zou *et al.*, 2006). SPOP also targets AR for degradation (An *et al.*, 2014), while mutant SPOP is unable to bind to, or mediate the degradation of, either SRC-3 or AR (An *et al.*, 2014; Geng *et al.*, 2013). Both AR and SRC-3 are necessary for androgen-dependent and androgen-independent cell proliferation in prostate cancer cell lines *in vitro* and xenografts *in vivo*, and are recruited to the promoters of cell cycle genes such as Cyclin A to promote their expression (Zou *et al.*, 2006). Another co-activator for AR, SRC-2 (alternatively known as NCOA2) is frequently amplified in prostate cancer, with up to 20% of primary tumours and 63% of metastases having amplifications of the chromosomal region of 8q which contains the gene (Taylor *et al.*, 2010). A third example of AR coactivator alterations is TRIM24. SPOP mutations stabilise TRIM24, which also co-activates expression of cell cycle and AR target genes with AR, is associated with tumour recurrence and is progressively overexpressed in prostate tumours and metastases (Groner *et al.*, 2016). While SPOP mutations have been found in high-grade PIN lesions suggesting that the event occurs early in tumorigenesis (Barbieri *et al.*, 2012), it is likely that it plays a greater role in the acquisition of castration resistance in this subset of tumours. The high burden of mutational events in SPOP and SRC-2 suggest another way of sustaining AR signalling: through mutation of its coactivators and regulators.

Overall, the acquisition of a castration-resistant phenotype requires sustained androgen receptor signalling. It is clear that this may occur in many different ways, including through AR amplification and mutation, but also through mutations in the regulators of AR signalling and cross-talk with other key oncogenic signalling pathways such as PI3K-AKT. This is reflected in the accumulation of mutations in these pathways throughout prostate tumorigenesis. The fact that key drivers of the PIN to carcinoma transition also appear to play roles in castration resistance might help to explain why tumours treated with hormone therapy invariably become castration resistant.

1.2.4 Summary

PIA is one of the hypothetical origins of prostate carcinoma, through PIN as the intermediate state (Figure 1.4). From PIA to PIN, the following are observed: a progressive reduction in the expression of antioxidant genes such as GSTP1 and NKX3-1; an increase in epithelial cell proliferation; expression of anti-apoptotic genes such as BCL2; and a progressive loss of basal cells and disruption of the glandular basement membrane. If PIA is the initial lesion, this would suggest that inflammation and the resulting oxidative stress are some of the first oncogenic insults in prostate cancer development, though this has yet to be proven and PIA's status as a precursor lesion is still uncertain.

The transition from PIN to carcinoma is marked by other key driver mutations including loss of PTEN and ETS gene fusion and SPOP mutation events which result in upregulation of PI3K-AKT and AR signalling respectively. PI3K and AR signalling are both key driver pathways of prostate cancer progression. Amplification of c-myc, upregulation of AMACR and loss of $p27^{Kip1}$ are also observed during this phase of progression, probably reflecting changes in metabolism, cell cycle progression and the escape of apoptosis in the carcinoma environment. Progression to advanced carcinoma likely coincides with p53 mutation or loss for further downregulation of apoptosis.

Transition from an advanced primary carcinoma to a castration resistant carcinoma is dependent upon selection induced by hormone therapy. In castration resistant prostate carcinoma, AR signalling persists in the absence of hormone and/or the presence of AR inhibitor. This event is mainly brought about by mutations or amplifications in the androgen receptor, and continued AR signalling is also sustained by aberrations in PI3K-AKT signalling and the pre-existing PTEN loss, ETS-fusions or SPOP mutations which alter these signalling pathways in the castrate environment.

1.3 Cancer Stem Cells (CSCs) in prostate cancer

This section will introduce the cancer stem cell hypothesis, a relatively new research area in prostate cancer, and critically review the current evidence for the existence of prostate cancer stem cells. It will also cover the phenotypes that are associated with prostate cancer stem cells, including self-renewal, multipotency and resistance to multiple therapeutic strategies. Some of the content in this section has been published by the author in a previous review article on methodologies for isolating and characterising cancer stem cells in the prostate (Sharpe *et al.*, 2013).

1.3.1 The Cancer Stem Cell Hypothesis

Stem cells are cell types that give rise to the terminally-differentiated, specialised cells of organs during development, and they continue to play a role in adult tissues in maintaining the homeostasis of their compartments by replenishing the differentiated cell populations found within (Blackwood *et al.*, 2011). The cancer stem cell (CSC) hypothesis postulates that tumour masses may arise from a single cancer cell with stem-like characteristics. These CSCs, sometimes known as "Tumour-Initiating Cells", are thought capable of self-renewal and differentiation to regenerate the tumour mass and all tumour cell types within it. Such cancer cells were first identified in leukaemia in the 1990s (Bonnet and Dick, 1997; Lapidot et al., 1994) followed by discoveries in breast cancer (Al-Hajj et al., 2003), and subsequently in other solid tumours (Visvader and Lindeman, 2008; Tirino et al., 2013) including prostate cancer (Collins et al., 2005). More recently, lineage-tracing evidence in the mouse suggests that putative prostate epithelial stem cells, such as the luminal castration-resistant NKX3-1-expressing cells known as CARNs, can act as a cell of origin for prostate carcinoma when given an oncogenic insult in the form of PTEN deletion (Wang *et al.*, 2009). This lends more evidence to the argument that stem-like cancer cells may be important in tumour development, but it is currently unclear whether a stem cell acts as a cell of origin in the context of human prostate carcinogenesis.

CSCs appear to be rare within tumours, as only a small proportion of all prostate cancer cells are able to reliably form large clones in vitro (Collins et al., 2005) and xenograft prostate tumours in mice *in vivo* (van den Hoogen et al., 2010; Hurt et al., 2008). Another feature attributable to cancer stem cells is their resistance to conventional treatment regimens, particularly chemotherapy (Abdullah and Chow, 2013) and the continued survival and self-renewal of CSCs is potentially an enticing explanation for

relapse, metastasis and therapy failure in prostate cancer. It is currently unclear whether CSCs are transformed stem cells or cancer cells that acquire stem-like phenotypes. A concerted effort is underway to isolate, characterise and target this critical cell population in order to produce more effective therapies.

Cancer stem cells provide a compelling theory with which to explain the heterogeneity within tumours, particularly in prostate tumours which have marked heterogeneity even within the same patient. Patients tend to have multiple foci of carcinoma which are genetically and phenotypically distinct (Abate-Shen and Shen, 2000), apparently due to separate epithelial cells acquiring mutations independently of one another (Cheng *et al.*, 1998; Blackwood *et al.*, 2011). Thus, it is possible to find benign tissue, PIN lesions and adenocarcinoma of multiple severities within the same biopsy and the same section (Abate-Shen and Shen, 2000), a heterogeneity already partly described and categorised by the Gleason scoring system.

Prostate carcinoma is also notoriously heterogeneous between individual patients. Transcriptomic analysis of individual patient tumours suggests that each tumour may bear its own diverse and distinct set of mutations, genomic rearrangements and novel transcripts (Wyatt *et al.*, 2014), and the molecular profiles are heterogeneous enough to confound attempts to construct a prognostic model that would provide extra predictive utility over standard clinical variables (Sboner *et al.*, 2010). Successful attempts have been made at molecularly subtyping the disease, resulting in seven tumour subtypes based on gene mutations (Abeshouse *et al.*, 2015), but even so over a quarter of all tumours remain unclassified. Thus, the disease presents a considerable challenge for the identification of reliable prognostic models that would help inform treatment decisions for patients with prostate cancer. Theoretically, identifying the cancer stem cells from which these heterogeneous populations emerge could provide useful insight into the tumour's development and subsequent disease progression. Identifying CSC-related biomarkers for use in a pathology setting might allow for a more accurate prediction of disease course and tumour recurrence.

1.3.2 The identity of prostate CSCs

The identity and the existence of prostate CSCs are both controversial, and there have been many attempts to isolate and characterise them, mainly from prostate cancer cell lines, but also from primary human tumour tissue. There are many potential biomarkers used to identify putative prostate CSCs, and this section will cover some of the frequently used protein biomarkers, such as CD44 and CD133, which are used to identify putative CSCs in several different tumour types including prostate, breast and gliomas (Visvader and Lindeman, 2008). Other CSC biomarkers will be reviewed in subsequent results chapters. These biomarkers are used either alone or in combination with other biomarkers to indicate the presence of putative cancer stem cells.

1.3.2.1 CD44 as a prostate CSC biomarker

CD44 is a hyaluronan-binding cell-surface glycoprotein that is often used to purify prostate cancer stem cells by fluorescence-activated cell sorting. CD44 has multiple splice isoforms and regulates many cellular processes including cell migration, proliferation and cell survival, which appear to be dependent on numerous binding partners such as Ezrin, Radixin and Moesin proteins and receptor tyrosine kinases such as c-Met (Ponta *et al.*, 2003). Evidence from cell lines and xenograft experiments, detailed below, supports potential roles for CD44 in mediating CSC behaviour.

SiRNA-mediated knockdown of splice variant CD44v6 has been found to reduce clonogenicity and sphere forming capacity in LNCaP, PC-3M and DU145 cells (Ni *et al.*, 2014). *In vitro*, orthotopic mouse xenografts of CD44+ subpopulations of prostate cancer cell lines DU145, PC-3 and LNCaP have greater proliferative, clonogenic and sphere-forming capacity than CD44- cells, as do immortalised normal human prostate epithelial NHP6 cells (Patrawala *et al.*, 2006). Interestingly, even non-tumorigenic NHP6 cells had a subpopulation of 2% CD44+ cells, although minimally tumorigenic LNCaP cells had none and tumorigenic PC-3 cells had 100%, which suggests that CD44 expression correlates with tumorigenic potential *in vivo*, as well as proliferation *in vitro*. When orthotopically xenografted into immunodeficient mice, these cell lines give rise to primary tumours and metastases more frequently, with as few as 100 implanted cells required for tumour formation (Patrawala *et al.*, 2006), whereas CD44- and unsorted cells require several orders of magnitude more cells for successful tumour formation. Conversely, orthotopic xenografting of LAPC4, DU145 and PC-3 cells with silenced CD44 expression results in reduced tumour growth and metastasis (Liu *et al.*, 2011a),

indicating that CD44+ tumour cells have a greater capacity for proliferation and invasion.

CD44+ CSCs from prostate cancer cell lines display evidence of stem-like properties: BrdU label-retaining properties indicate a slower cycling time, and they also express increased levels of stem cell markers such as BMI1, SMO, OCT3/4 and beta-catenin (Patrawala *et al.*, 2006). However, CD44- can give rise to CD44+ cells in prostate cancer cell lines, suggesting that this may be a dynamically regulated or fluctuating phenotype (Patrawala *et al.*, 2006).

However, the effect of CD44 expression on prostate tumours is far from clear: one study showed that CD44s and splice variant CD44v6 protein expression may be lost in metastases, and expression of CD44s inversely correlates with Gleason grade (De Marzo *et al.*, 1998). Overexpression of CD44 in a rat prostate cancer cell line suppressed metastasis formation in subcutaneous murine xenografts (Gao *et al.*, 1997). In summary, it is recognised that human prostate cancer cell lines are not homogeneous populations in terms of their proliferative and invasive capacities, but rather a stem-like CD44+ population has elevated tumorigenic potential, although this remains controversial. However, the putative CD44+ prostate CSC population is thought to be highly heterogeneous, and therefore when characterizing prostate cancer stem cells CD44 is often included in a panel of biomarkers of prostate CSCs (Table 1.1).

Table 1.1: A summary of prostate cancer stem cell marker panels identified in the literature.

Expression Profile	*Cells/Tissues Studied*	*Evidence for Cancer Stem Cell Identity*	*Source*
$CD44^{+}/CD24^{-}$	Prostate Cancer Cell Lines, Prostate Cancer Xenografts	Anchorage-independent growth; Increased clonogenicity; Prostatosphere Formation; Expression of stem cell markers; Tumour formation in mice at low titres.	(Hurt *et al.*, 2008)
$CD44^{+}/CD24^{+}/ \alpha_2\beta_1^{high}$	Prostate Cancer Cell Lines, Prostate Cancer Xenografts	Prostatosphere formation; Serial prostatosphere passage; Tumour formation in mice at low titres.	(Rybak *et al.*, 2011)
$CD44^{+}/CD133^{+}/ \alpha_2\beta_1^{high}$	Explanted Primary Human Prostate Tumour	Anchorage-independent growth; Increased clonogenicity; Cells differentiate upon culture with serum.	(Collins *et al.*, 2005)
$CD44^{+}/CD133^{+}/ ABCG2^{+}/CD24^{-}$	Explanted Primary Human Prostate Tumour	High clonogenic potential; Prostatosphere formation; Immunohistochemistry for CD133, CD44 and ABCG2 in prostate cancer biopsies.	(Castellon *et al.*, 2012)
$PSA^{-/low}/ALDH^{+}/ CD44^{+}/ \alpha_2\beta_1^{+}$	Prostate Cancer Cell Lines, Prostate Cancer Xenografts	Tumour formation in mice at low titres. Increased cell frequency following castration.	(Qin *et al.*, 2012)

1.3.2.2 Other biomarkers for prostate CSCs

Other cancer stem cell markers are frequently co-expressed in CD44+ prostate CSCs. $CD44^{+}/CD24^{-}$ purified DU145 (Salvatori *et al.*, 2012) and LNCaP (Hurt *et al.*, 2008) prostate cancer cell lines form tumourspheres with the potential to differentiate. Stem-like $PSA^{-/low}$ cells sorted for the antigenic profile $ALDH^{+}/CD44^{+}/ \alpha_2\beta_1^{+}$ had higher tumorigenicity than $ALDH^{-}/CD44^{-}/ \alpha_2\beta_1^{-}$ cells and PSA^{-} cells, with ten cells being sufficient to induce a xenograft tumour in NOD/SCID mice (Qin *et al.*, 2012). Similarly, tumorigenic prostate cancer stem cells, comprising 0.1% of the total cells, have been isolated from primary human prostate tumours using the antigenic profile $CD44^{+}/\alpha_2\beta_1^{high}/CD133^{+}$ (Collins *et al.*, 2005). CD133 is a 5-pass transmembrane glycoprotein expressed by some normal human stem cell populations, and was first identified in haematopoietic stem and progenitor cells (Yin *et al.*, 1997). Interestingly, CD133 is also expressed in a small subset of quiescent, $\alpha_2\beta_1^{high}$ prostate basal cells in benign human prostate tissue that have the capability of forming differentiated acinar structures, suggesting that populations of stem-like cells in the benign and malignant prostate epithelium may be related (Richardson *et al.*, 2004). CD133+ subpopulations of prostate cancer cell lines CWR22RV1, C4-2, LNCaP have greater invasion and sphere-forming capacity *in vitro* than the corresponding unsorted lines (Wen *et al.*, 2016). Again, this suggests that prostate epithelial cells expressing stem cell markers have more invasive phenotypes, and that tumour cells may be remarkably heterogeneous, even in cell lines that were previously thought to be relatively homogeneous populations. If these findings translate into the human tumour, then they carry prognostic implications for the patients with higher burdens of cells with these stem-like and CSC-like expression patterns.

Biomarker panels for prostate CSCs remain controversial. In CSCs isolated from cancer cell lines, such as DU145 tumoursphere cells (Rybak *et al.*, 2011), the prostate CSC population was $CD44^{+}/CD24^{+}/ \alpha_2\beta_1^{+}$, in contrast to the $CD44^{+}/CD24^{-}$ profile mentioned previously. Similarly in head and neck squamous cell carcinoma cell lines, $CD44^{-}$ cells had CSC-like properties despite its use as a marker for CSCs (Oh *et al.*, 2013). In IHC experiments conducted on prostate tumours, high-grade PIN and normal prostate tissue, CD44 was found to be more frequently expressed in PIN and normal tissue than in carcinoma, but in CD44+ cases the expression is too widespread to represent a rare stem-like cell population (Ugolkov *et al.*, 2011). This suggests that there may be more than one group of stem-like cells to be found in cancer cell populations, and a single

antigen - or panel of antigens - may be insufficient for isolating all CSCs of interest. Additionally, these CSCs may not necessarily be cells of origin for prostate cancer: in an *in vivo* murine renal capsule xenograft model of prostate cancer, only $CD133^-$ basal-enriched human prostate epithelial cells derived from the benign prostatic hyperplasia cell line BPH1 were vulnerable to transformation, either by estrogen and androgen administration or by recombination with cancer-associated fibroblasts (Taylor *et al.*, 2012). Despite their stem cell marker expression and functional capacity for proliferation and invasion, it is still unclear whether these prostate CSCs, mostly from prostate cancer cell lines, represent the cells of origin of the tumour, or whether they are tumour cells that have activated stem-like programmes. The hypothesis is currently supported by evidence drawn from cancer cell lines and their xenografts, which are repeatedly passaged and not representative of the human tumour microenvironment. While CD44 and CD133 are commonly used biomarkers for prostate CSCs, no single marker or combination of markers is sufficient for identifying all putative prostate CSCs, and it remains to be seen whether such a biomarker exists.

1.3.3 Functional assays for prostate CSCs

There have been many functional assays employed to identify and characterise prostate CSCs, each with a set of advantages and limitations. These will be referred to in the next section on CSC properties and subsequently in chapter 3, and so will be discussed here, separated by the CSC properties they characterise.

1.3.3.1 Self-renewal

A defining property of both stem cells and cancer stem cells is their self-renewal capacity. There are two main types of in vitro assays used to probe self-renewal in isolated cell populations: clonogenicity assays and sphere formation assays. Both assays involve the culture of isolated cells in conditions that preferentially maintain stem cells, thus selecting for those cancer cells with more stem-like properties *in vitro*. They may also be considered proliferation assays, as self-renewal and proliferation are not distinct in this context. As most cancer cells are expected to have some self-renewal capacity, these assays are comparative in nature, and rely on the isolated CSCs being significantly different from a control population of cells for a tested phenotype.

1.3.3.1.1 Clonogenicity assay

The clonogenicity assay involves seeding a single cell or a small number of cells onto culture plates, and monitoring colony formation after a defined time period. This assay is based on the assumption that small titres of cells will only form large colonies if they have the ability to self-renew (Hellsten *et al.*, 2011). The number of colonies present is then treated as a correlate of self-renewal capacity. This assay can be done serially with a single cultured cell, making it possible to further demonstrate a clone's self-renewal capacity over multiple passages *in vitro* (Collins *et al.*, 2005). There are some disadvantages with this assay; one being that an arbitrary threshold must be assigned to determine which colonies are too small to count. More significantly the assay may not provide a complete representation of self-renewal capacity *in vivo*. However, they underline the important finding that subpopulations of stem-like prostate cancer cells have far greater proliferative capacity than the rest of the cell population.

1.3.3.1.2 Sphere formation assay

Another important tool is the sphere formation assay. The method was originally used to study adult neural stem cells *in vitro* (Reynolds and Weiss, 1992) and has now been applied to isolate prostate CSCs. Low adherence culture in defined serum-free medium produces clonal multicellular spheroid aggregates called tumourspheres, which are

enriched for cells displaying cancer stem cell phenotypes. Anchorage-independent growth is a malignant phenotype in itself, suggesting that isolated CSCs are selected by their innately malignant phenotypes. This method is now widely-used for the enrichment of cancer stem cells from many types of cancer, including prostate cancer. The assay is versatile, and may be used in prostate cancer cell lines (Zhang *et al.*, 2012a; Cho *et al.*, 2012; Miki *et al.*, 2007) and explanted human primary prostate tumour tissue (Rajasekhar *et al.*, 2011; Miki *et al.*, 2007; Castellon *et al.*, 2012), or to further characterize cells isolated by other methods (Celia-Terrassa *et al.*, 2012; Albino *et al.*, 2012; Ma *et al.*, 2011; Hellsten *et al.*, 2011). CD44 was found to be highly expressed in tumourspheres derived from four different prostate cancer cell lines (Zhang *et al.*, 2012a), suggesting that expression of this marker may be important for the anchorage-independent growth phenotype.

Tumourspheres form at low efficiency, originating from a small fraction of the tumour cell population (Rajasekhar *et al.*, 2011) and can be serially cultured (Doherty *et al.*, 2011; Duhagon *et al.*, 2010; Rajasekhar *et al.*, 2011; Kong *et al.*, 2010), which further indicates that prostate cancer cells and their tumoursphere progeny both contain a subpopulation with extended proliferative capacity. Tumoursphere cells also tend to possess CSC-associated phenotypes, such as self-renewal in long-term culture and expression of putative cancer stem cell markers such as CD44 and integrin $\alpha_2\beta_1$ (Rybak *et al.*, 2011; Zhang *et al.*, 2012b). However, sphere formation may not always select for cancer stem cells *in vitro*: $\alpha_2\beta_1^+$ prostate cancer stem cells isolated from prostate xenograft tumours had high clonogenicity and sphere-forming capacity, but did not have the high tumorigenicity in xenografts expected of cancer stem cells (Patrawala *et al.*, 2007). Similarly, when Matilainen *et al.* (2012) attempted to isolate breast cancer stem cells using this method, the result was an unexpected reduction in CSC-associated gene expression and a loss of metastatic phenotypes. Tumourspheres may not consist purely of CSCs, but contain some differentiated and dying cells (Salvatori *et al.*, 2012), and there are concerns that sphere generation is not being sufficiently validated – for example, to ensure that sphere formation is a result of cell proliferation and not cell aggregation (Pastrana *et al.*, 2011). Overall, these results suggest that tumoursphere formation and thus anchorage-independent growth is a malignant phenotype possessed by prostate CSCs *in vitro*. It is not currently known whether this is a property of CSCs in a tumour microenvironment *in vivo*, highlighting the need to verify all findings *in vivo* before any conclusions are made regarding the existence of prostate CSCs.

1.3.3.1.3 Xenograft transplantation assays

In vitro assays have limited utility in characterising cancer stem cells, as they take isolated epithelial cells out of the context of their tumour microenvironment. A well-established test for *in vivo* self-renewal comes from xenograft transplantation assays. Low titres of CSCs – isolated from cell lines (van den Hoogen *et al.*, 2010) or from dissociated human primary (van den Hoogen *et al.*, 2010) and xenograft (Rajasekhar *et al.*, 2011) prostate tumours are injected into immunodeficient mice. Prostate CSCs with a variety of markers such as high aldehyde dehydrogenase activity, primary tumours (Cho *et al.*, 2012; van den Hoogen *et al.*, 2010; Rajasekhar *et al.*, 2011; Li *et al.*, 2010) or metastases (van den Hoogen *et al.*, 2010) at this limiting dilution are proposed to have a cancer stem cell phenotype. To provide more evidence for the CSC identity of the cells, xenograft tumour formation assays can be done serially using purified dissociated cells from previous xenograft tumours (Li *et al.*, 2010). The ability to serially generate tumours at limiting dilution is an indicator of self-renewal *in vivo*. Large tumour size and high tumour-forming capacity are indicators of an aggressive subpopulation of cells *in vivo*. Although this does not fully represent tumour formation *in situ*, especially as the adaptive and innate immune systems are compromised, these models are versatile and demonstrate tumorigenicity and proliferative capacity *in vivo*. In fact, this approach was perhaps the gold standard model for studying CSCs, until the advent of *in vivo* lineage tracing approaches (see Chapter 1.3.3.2.3).

1.3.3.2 Multipotency

1.3.3.2.1 Matrigel assays

Another defining characteristic of normal stem cells and, hypothetically, cancer stem cells, is their ability to differentiate. Differentiation can be induced *in vitro* by culturing cells in a scaffold of matrigel, and multipotent prostate cancer cells may produce glandular structures that recapitulate prostate cancer architecture *in vivo* (Gu *et al.*, 2007). Unlike culture in serum-free media, matrigel induces differentiation, causing putative prostate stem cells to form glandular structures. CSCs may also differentiate in standard mammalian cell culture with added serum and dihydrotestosterone (Collins *et al.*, 2005; Hurt *et al.*, 2008), losing stem cell markers and concurrently expressing differentiated prostate epithelial cell proteins such as AR and prostatic acid phosphatase (Hurt *et al.*, 2008). This assay demonstrates that prostate CSCs may assume more

specialised cell types when exposed to the appropriate differentiation factors, recapitulating another of the characteristic stem cell properties *in vivo*.

1.3.3.2.2 Tissue recombination assays

To provide more evidence for differentiation capacity *in vivo*, prostate CSCs can be recombined with inductive rat urogenital sinus mesenchyme and grafted under the renal capsule of immunodeficient mice. With a stromal component provided, gland-like structures form that recapitulate the structures and cell types found in the original prostate (Goldstein *et al.*, 2010) or human prostate tumour (Gu *et al.*, 2007), thus demonstrating a stem-like capacity for differentiation. The renal graft model can also be used to assay for malignant transformation capacity, by recombining selected prostate epithelial cells with cancer-associated fibroblasts from human prostate tumours (Taylor *et al.*, 2012). This assay highlights the importance of the stromal microenvironment for differentiation and cell transformation, which appears to be as relevant for prostate CSCs as normal stem cells. While these assays do provide a stromal component which better recapitulates the *in vivo* situation, the removal of putative prostate CSCs from their original tumour context may create a false impression of stemness behaviour that does not truly exist in the tumour proper: stem cells are highly dependent on their stem cell niche (Foster *et al.*, 2002). For example, AR expression in the prostatic mesenchyme is necessary for the prostate gland and its secretory epithelium to form (Kurita *et al.*, 2001). As with normal stem cells, the existence of prostate CSCs cannot be proven until such cell behaviours are observed in tumours *in situ* and *in vivo* – something that is very difficult to achieve in human tumours. To better address this question, *in vivo* lineage tracing studies have been employed in the mouse to model the development of prostate cancer.

1.3.3.2.3 In vivo lineage-tracing

The presence of prostate CSCs in human prostate cancer remains to be proven. However, the lineage-tracing technique, which irreversibly marks a cell population and its progeny for long-term tracing of cell fate, was described in Chapter 1.1.3, and has been used in mouse models to monitor prostate epithelial cells following an oncogenic insult. Lineage-tracing evidence is limited for the prostate cancer, but some studies have been conducted to probe potential cells of tumour origin. Luminal castration-resistant NKX3-1-expressing cells, which are a putative luminal stem/progenitor cell population in the developing prostate, can act as a cell of origin for prostate carcinoma when PTEN

is deleted in a model of prostate regression and regeneration (Wang *et al.*, 2009), resulting in a tumour with a luminal phenotype and loss of basal cells. These alterations recapitulate those seen in human prostate tumours as detailed in Chapter 1.2.3. A potential criticism is that most mouse models of prostate cancer already make use of a luminal-specific promoter to drive expression of transgenes, such as the probasin promoter or its stronger derivative ARR2PB (Majumder *et al.*, 2008), and therefore the finding of tumorigenesis from this compartment may not be novel. However, a more recent lineage-tracing study found that stabilisation of beta-catenin in Axin2-expressing luminal cells, which were demonstrated to be unipotent luminal progenitors in the mouse prostate during development, resulted in adenocarcinoma in situ with luminal differentiation and expression of proliferative marker Ki-67 (Lee *et al.*, 2015b). The results of lineage-tracing experiments so far suggest that transformation in putative luminal stem/progenitor compartments can result in prostate carcinoma in mice.

Interestingly, basal cells can also be the targets for oncogenic transformation. The same group that generated NKX3-1-driven PTEN-null mouse tumours later demonstrated that luminal phenotype prostate tumours could also be generated from the basally expressed promoter CK5 using the same PTEN deletion as an oncogenic insult (Wang *et al.*, 2013). Basal-origin tumours progress slower to carcinoma than luminal-origin tumours, but they are histologically similar at the equivalent stages (Wang *et al.*, 2013), perhaps suggesting that the basal transformation process to carcinoma takes a longer time due to a preceding basal-to-luminal differentiation event.

While stem cell compartments can be transformed into carcinoma in mouse lineage-tracing models, it remains to be seen whether tumour-forming cell populations in the mouse prostate continue to act as prostate CSCs during development of the disease, or whether they only contribute to initiation in this model. These cells satisfy some of the requirements for being prostate CSCs in that they have considerable self-renewal as well as apparent differentiation potential. However, the contribution of Axin2, CK5 or NKX3-1-expressing tumour cells to recurrence, therapy resistance or metastasis, which are also functionally important aspects of CSC behaviour, has yet to be demonstrated.

1.3.4 Other properties of prostate CSCs

1.3.4.1 Expression of stem cell markers

As mentioned in the previous section, some prostate CSC markers are also stem cell markers, and a recapitulation of stem-like expression patterns and behaviours is one of the hallmark characteristics of a CSC. Many embryonic stem (ES) cell pluripotency factors – such as Nanog, Oct4 and Sox2 - are known to be expressed in both adult stem cells and CSCs, and embryonic stem cell gene expression signatures have been found to identify poorly differentiated tumours from several types of human cancer including prostate cancer (Ben-Porath *et al.*, 2008), and this has been verified in subsequent studies (Rye *et al.*, 2014; Markert *et al.*, 2011). Given this trend, stem cell factors likely underlie many of the phenotypes observed in prostate CSCs, as they may underpin important cancer stem cell phenotypes such as self-renewal. DU145 prostate tumoursphere cells express the ES pluripotency factors Nanog, Oct4 and Sox2 (Rybak *et al.*, 2011), as do telomerase-immortalized human prostate cancer epithelial cell cultures (Gu *et al.*, 2007). In human BM18 xenograft prostate tumours, rare $CD44^+$ stem-like cells have also been found to express Nanog (Germann *et al.*, 2012). DU145 and PC-3 cells that express GFP under a Nanog promoter reporter construct have been shown to have increased clonogenicity and tumoursphere formation capacity, as well as having forming larger tumours in murine xenografts (Zhang *et al.*, 2014). Interestingly, in both the xenograft and tumoursphere settings, Nanog promoter activation was present in tumours or spheres that were derived from GFP-negative cells, suggesting dynamic regulation of Nanog expression in the putative prostate CSC population (Zhang *et al.*, 2014). The functional role of Nanog expression in these cells appears to be cell cycle progression, as GFP-positive cells had greater levels of CDK2, cyclin E, and cyclin A protein than GFP-negative cells, suggesting that Nanog expression promotes entry into S-phase (Zhang *et al.*, 2014). In primary human prostate carcinoma, expression of Nanog and Oct4 has been reported to be enriched in Gleason grade 5, an indicator of poor prognosis (Mathieu *et al.*, 2011). This evidence suggests that the presence of stem cell markers may be a common feature of poor-prognosis tumours.

However, the association between pluripotency factors and CSC properties remains controversial and incompletely understood, especially in patient tumours. Sox2 expression was recently found to be repressed by signalling through the Androgen Receptor in prostate cancer cell lines (Kregel *et al.*, 2013), and was associated with castration-resistant prostate carcinoma tissues assessed by IHC. Expression in human

prostate tumours was not consistent with a CSC population, as it was broadly expressed throughout Sox2-positive tumour epithelium (Kregel *et al.*, 2013). Conversely, an IHC study showed that Sox2 was expressed in both benign prostate and prostate cancer in a small subpopulation of <10% of cells, which is more consistent with a stem cell population.

In addition, both Oct4 (Sotomayor *et al.*, 2009) and Nanog (Das *et al.*, 2011; Booth and Holland, 2004) have multiple pseudogenes and splice variants which could have differing importance to prostate CSCs and thus complicate the interpretation of the overall roles of these proteins in these cell populations. For example, in patients with primary prostate carcinoma, the Oct4 isoform Oct4A is expressed in the nuclear compartment and is associated with a shorter biochemical recurrence-free survival, whereas Oct4B is expressed in the cytoplasm and associated with a longer biochemical recurrence-free survival (de Resende *et al.*, 2013), suggesting different roles for the two isoforms of the protein.

Nanog mRNA is expressed mainly through the functional pseudogene NANOGP8 in prostate cancer cell lines, including LNCaP, DU145 and PC-3 (Jeter *et al.*, 2011; Zhang *et al.*, 2014), and Nanog protein levels positively correlate with CD44 protein levels (Jeter *et al.*, 2011). In HPCa38 and HPCa39 prostate cancer xenografts and in PC-3 and DU145 cells, shRNA-mediated knockdown of Nanog levels results in inhibition of proliferation and clonogenicity, whereas knockdown of Oct4 does not have a clear inhibitory effect on proliferation (Jeter *et al.*, 2011). In summary, while there seems to be a more consistent role for Nanog expression in prostate CSCs, the expression pattern and significance of Oct4 and Sox2 expression in both prostate tumours and their putative CSCs is still unclear, but it seems unlikely that prostate CSCs are a simple recapitulation of embryonic stem cell transcription in cancer cells.

1.3.4.2 Expression of prostate epithelial cell lineage markers

Markers that identify specific cell types of the prostate are often used in attempt to characterise candidate prostate cancer stem cells. This is important, as the normal prostate cell type that prostate cancer stem cells most resemble remains controversial. For example, populations of castration resistant stem-like cells persist in DuCaP and HPET cell lines *in vitro* (Pfeiffer *et al.*, 2011; Gu *et al.*, 2007) and in prostate cancer xenografts *in vivo* (Germann *et al.*, 2012). Germann et al. have suggested that the CSCs found in the BMI18 xenograft model are of luminal or possibly neuroendocrine cell

origin, detecting elevated expression of the luminal marker Cytokeratin 18 and neuroendocrine cell marker Chromogranin A following castration (Germann *et al.*, 2012). Conversely, another group observed upregulation of basal-specific cytokeratin 5 in the DuCaP cell line under androgen deprivation, but could not detect expression of the neuroendocrine marker Nestin (Pfeiffer *et al.*, 2011). However, when comparing expression profiles of castration resistant and non-resistant human prostate tumours by microarray, the group found that both basal (Cytokeratin-5) and neuroendocrine (Nestin) markers were upregulated along with the putative cancer stem cell markers CD44 and c-Kit. Immortalised prostate cancer epithelial cells generated by Gu *et al.* (2007) also express Nestin along with embryonic stem cell markers, consistent with a neuroendocrine-like CSC phenotype. Markers of prostate lineage differentiation can also be absent: docetaxel-resistant 22Rv1 and DU145 CSCs are negative for both luminal and basal cell cytokeratins as well as androgen receptor (Domingo-Domenech *et al.*, 2012). The variability of expression patterns in multiple models of prostate cancer suggests that cancer stem cells may be more than one distinct subpopulation. It could also be a reflection of the phenotypic plasticity of cancer stem cells that has been recently modelled in breast cancer cell lines *in vitro*, where cancer cells reach an equilibrium of available differentiation states even once purified for a specific differentiation state (Gupta *et al.*, 2011; Chaffer *et al.*, 2011) and in xenografts of these lines *in vivo* (Chaffer *et al.*, 2011). Potential examples of this in the prostate include spontaneous interconversion of CD44+ and CD44- cell states in LAPC-4 and LAPC-9 prostate cell line xenografts (Patrawala *et al.*, 2006), and interconversion of Nanog+ and Nanog- cell states in PC-3 and DU145 xenografts (Zhang *et al.*, 2014). Regardless, determining the cell type of origin – and the heterogeneity thereof - is important in order to further establish the identity of prostate cancer stem cells.

Overall, the epithelial cell type that putative prostate CSCs most resemble is still unclear. Biomarkers are useful for probing all aspects of cancer stem cells but given the apparent heterogeneous nature of cancer stem cells, currently no single marker is ideal and the best current practice is to use combinations of markers analysed in tandem for effective phenotypic characterization.

1.3.4.3 Regulation by microRNAs

Cancer stem cell phenotypes appear to be modulated in part by microRNA-mediated mechanisms which contribute to the expression of stem cell markers in prostate CSCs.

miR-34a, a microRNA that is a transcriptional target of p53, downregulates CD44 expression and is consequently downregulated in CD44+ and CD133+ prostate cancer stem cells from three xenograft models (Liu *et al.*, 2011a). In prostate cancer cell lines, miR-34a expression results in reduced clonogenicity and sphere formation *in vitro* as well as reduced tumour growth and metastasis when implanted into mice *in vivo* (Liu *et al.*, 2011a). This appears to be largely due to suppression of CD44 expression, although expression of miR-34a *in vitro* can also lead to apoptosis and cell cycle arrest in G_1-phase (Tarasov *et al.*, 2007). Administration of miR-34a may thus have a potential therapeutic use, and highlights the clinical potential inherent in the inhibition of CSC phenotypes.

Like miRNA-34a, Let-7 is upregulated by p53 expression (Tarasov *et al.*, 2007) and is responsible for the silencing of oncogenes c-MYC (Liu *et al.*, 2012), k-RAS (Johnson *et al.*, 2005; Liu *et al.*, 2012) and HMGA2 (Lee and Dutta, 2007), which enhance cell proliferation. Let-7 expression also silences the stem cell markers OCT3/4, SOX2 and NANOG in PC-3 cells *in vitro*, and Let-7 levels themselves are reduced by expression of an RNA binding protein known as Lin28B (Kong *et al.*, 2010). Lin28B is inhibited by expression of another microRNA, miR-200 (Kong *et al.*, 2010), whose expression in PC-3 cells that have an EMT-like phenotype causes them to undergo MET and have reduced clonogenic and sphere-forming capacity in vitro. Lin28B also reduces expression of NOTCH1, another stem cell marker (Kong *et al.*, 2010). Lin28B is implicated in the acquisition of CSC phenotypes and pluripotency gene expression through suppressing Let-7 in oral squamous cell carcinoma cells (Chien *et al.*, 2015) and in ALDH activity and pluripotency gene expression in a variety of other carcinoma cell lines (Yang *et al.*, 2010), suggesting that the regulation of CSC phenotypes by this miRNA network may be independent of tumour origin.

Interestingly, ETS factor EHF, which has previously been shown to repress EMT and pluripotency-related genes as well as being reduced in prostate cancer (Albino *et al.*, 2012), regulates the expression of LIN28B by transcriptional repression of its promoter (Albino *et al.*, 2016). Thus, a network of microRNAs and their regulators is involved in the regulation of stemness and EMT phenotypes in prostate cancer, which appear to coordinate with one another in a cancer stem cell context. In general, many tumour-suppressing miRNAs are commonly downregulated in most prostate CSCs to promote these phenotypes, but the patterns of miRNA regulation appear to be distinct in CD44+, CD133+, integrin $\alpha_2\beta_1^{high}$ and side-population cells (Liu *et al.*, 2012). This would

suggest that, while all of these markers enrich for CSC phenotypes, the CSC population is heterogeneous and each of these subpopulations may have different properties. With further characterisation, these microRNA expression profiles could be a useful method for profiling cancer stem cell phenotypes *in vitro*.

1.3.4.4 Resistance to Therapy

The segregation of cancer stem cells from a cell population may be achieved by selecting for their resistance to standard cancer treatments. This is based on the observation that cancer stem cell populations tend to survive exposure to chemotherapeutic agents *in vitro*. Similarly, cancer stem cells have been proposed to be resistant to radiation and androgen withdrawal.

1.3.4.4.1 Castration resistance

In an androgen-sensitive human xenograft model, Germann *et al.* have recently identified subpopulations of cancer cells that repopulate the tumour following a cycle of androgen deprivation and replacement (Germann *et al.*, 2012), implying a potential role for cancer stem cells in the development of castration-resistant prostate cancer. Miki et al. have shown that inducing differentiation in culture can cause stem-like AR^- tumoursphere cells to become AR^+ (Miki *et al.*, 2007), suggesting that lack of AR expression may be an important cancer stem cell phenotype. The anti-androgens bicalutamide and enzalutamide have the effect of amplifying the CD133+ CSC population and sphere-forming capacity in prostate cancer cell lines (Wen *et al.*, 2016), and it is tempting to speculate that hormone therapy could bring about amplifications in, or select for, the CSC population in patient tumours which might contribute to castration resistance. Testing for androgen sensitivity and AR expression are important characterization steps for any prostate cancer stem cell population.

1.3.4.4.2 Chemoresistance

It is thought that CSCs within prostate tumours will be more resistant to chemotherapy, thus resulting in a subpopulation that could contribute to recurrence. CD44+ CSCs have been found to have slow cycling times *in vitro* (Patrawala *et al.*, 2006), a phenotype which could hypothetically contribute to their ability to survive cytotoxic chemotherapy. There are several lines of evidence that support some prostate CSCs being chemoresistant.

Firstly, potentially chemoresistant cells exist in the side population, where live cells are stained with lipophilic dyes such as Hoechst 33342. Stem-like cells have a greater capacity for efflux of the dye, allowing for their separation from the rest of the population by cell sorting (Hirschmann-Jax *et al.*, 2004). Side populations that efflux lipophilic dyes have been identified in both benign prostate tissue and prostate tumours, comprising a subpopulation of less than 5% of cells in all cases (Mathew *et al.*, 2009; Brown *et al.*, 2007). However, only LAPC-9 xenograft tumours were found to have a detectable side population in one study (Patrawala *et al.*, 2005) and while CWR-R1, DU-145 and RWPE-1 cell lines had a detectable side population, BPH-1, LAPC-4 or PC-3 did not (Mathew *et al.*, 2009). Considering that cells with CSC-like phenotypes have been observed in many of these cell lines, this suggests that either the side population is not a distinguishing feature of all prostate CSCs, or that the side population of some cell lines is too small to detect.

Secondly, the ability to efflux lipophilic dyes is linked to the expression of ABC transporters such as ABCG2 (Zhou *et al.*, 2001; Hirschmann-Jax *et al.*, 2004) and ABCA3 and is a property shared with stem cell populations in the bone marrow (Hirschmann-Jax *et al.*, 2004). ABC transporter expression may be involved in the efflux of chemotherapeutic drugs such as mitoxantrone in a variety of cancer cell lines from many tumour types (Hirschmann-Jax *et al.*, 2004), including in prostate cancer cell line CWR-R1, where ABCG2 knockdown sensitises cells to mitoxantrone, docetaxel and toptecan (Xie *et al.*, 2008). Expressing ABCG2 in LNCaP cells also reduces the sensitivity to these drugs. The profound effect of ABCG2 expression on drug resistance *in vitro* potentially means that this cell population is therapeutically. However, the expression of ABCG2 may have other roles outside of drug resistance: while side population cells from LAPC-9 xenograft tumours are at least 100 times more tumorigenic in murine xenografts than their non-side population counterparts, this is not due solely to ABCG2 expression, which is present only in a minority of side population cells (Patrawala *et al.*, 2005). $ABCG2^-$ and $ABCG2^+$ prostate cancer cell line subpopulations can interconvert and are similarly tumorigenic (Patrawala *et al.*, 2005). Patrawala and colleagues propose that the $ABCG2^+$ population actually marks highly proliferative progenitor cells, while the $ABCG2^-$ cells represent cancer stem cells, based on the finding that many more ABCG2+ cells are mitotic.

Thirdly, drug resistant cells are enriched in prostate CSC and stem cell markers. Recently, ABCG2 has been found to be expressed along with a panel of prostate CSC

biomarkers ($CD133^+/CD44^+/ABCG2^+/CD24^-$) in tumoursphere cultures from human prostate cancer explants (Castellon *et al.*, 2012). CSC markers may have functional relevance in the context of drug resistance: when CD44 splice variant CD44v6 is knocked down *in vitro,* prostate cancer cell lines LNCaP, PC-3M and DU145 are sensitised to a variety of chemotherapeutic drugs including paclitaxel, docetaxel, doxorubicin and methotrexate (Ni *et al.*, 2014). Conversely, docetaxel-resistant cell lines derived from the 22Rv1 and DU145 cell lines show decreased expression of both epithelial cytokeratins and androgen receptor, as well as upregulation of CD44 and the Notch and Hedgehog signalling pathways (Domingo-Domenech *et al.*, 2012). Hedgehog signalling has been described to mediate chemoresistance, EMT and CSC properties in several non-prostate cancer cell lines *in vitro* (Kong *et al.*, 2015). This would suggest that dedifferentiation and the acquisition of a CSC phenotype are important for the development of chemoresistance.

1.3.4.4.3 Radiation resistance

Some isolated cancer stem cells show resistance to ionizing radiation treatment. Cancer stem cells have been found to be radiation resistant in breast cancer cell lines (Yin and Glass, 2011; Fu *et al.*, 2012) and glioblastoma (Bao *et al.*, 2006). Currently this phenotype has not been widely characterized in prostate cancer stem cells. In the few prostate cancer cell lines that have been investigated, the radiation responses are variable (Cho *et al.*, 2012), with LNCaP tumoursphere-derived CSCs showing increased survival and reduced DNA damage compared to total cells, and DU145 CSCs showing no difference. However, the survival of CSCs in this assay was only apparent after a week of recovery, suggesting a period of DNA repair or adaptation is required. This phenotype may in part be mediated by CD44: knockdown of CD44 splice variant CD44v6 in prostate cancer cell lines LNCaP, PC-3M and DU145 sensitises them to a single dose of 4Gy of radiation (Ni *et al.*, 2014), reducing colony formation *in vitro*. Consistent with this, in human primary prostate tissue CD44 protein expression is elevated following radiotherapy (Cho *et al.*, 2012). Although these studies suggest a possible role for prostate CSCs in resistance to radiation, the mechanisms for this are unknown, and more work is required to determine whether radiation resistance is a distinguishing characteristic of prostate cancer stem cells.

1.3.5 Summary

The cancer stem cell hypothesis recognises that tumours are not homogeneous, but instead contain subpopulations of tumour cells with different properties. CSCs have been shown to recapitulate many stem-like properties in a cancer context, including expression of stem cell markers, the ability to self-renew and 'differentiate' to give rise to tumour cells with different properties, and therefore repopulate a tumour. CSC properties appear to be regulated by stem cell programmes, along with a network of microRNAs and pluripotency factors. Perturbing prostate CSCs in cell line and animal models has shown that many CSCs also survive common therapies for prostate cancer and therefore are critical may be causative agents of recurrence. However, this aspect of prostate CSCs requires much more characterisation. Additionally, it is currently unclear what prostate epithelial cell type prostate CSCs resemble, and in vivo lineage-tracing studies have found that multiple prostate epithelial cell types may act as cells of origin for prostate carcinoma, which may not necessarily be CSCs.

There is evidence to suggest that the properties of CSCs predispose towards a more malignant phenotype *in vivo*, although there are still no consensus markers or definitions for this heterogeneous cell population. If prostate cancer cells can switch stochastically between differentiation states, as has been shown in breast cancer cells and may be apparent in some prostate CSCs, it raises the possibility that non-stem cancer cells can give rise to stem-like cancer cells, which further complicates the definition of a CSC. Although the presence of CSCs in human tumours remains to be proven, the phenotypic characteristics of prostate CSCs identified so far make them a lucrative target for the development of new therapies and new biomarkers for prostate cancer.

1.4 Aims of this thesis

The goal of the thesis is to identify new prognostic biomarkers for predicting recurrence in patients with prostate cancer, using what has been learned about cancer stem cells to guide the search. There are three main aims:

i. To identify potential prognostic biomarkers relating to cancer stem cells or recurrence using published literature and bioinformatic approaches;

ii. To test these potential biomarkers in prostate carcinoma specimens, determining their relationship to recurrence and other clinical parameters;

iii. To further investigate and characterise any candidates generated through this approach in prostate carcinoma tissue.

2 MATERIALS AND METHODS

2.1 Materials

All general laboratory reagents were of molecular biology grade and purchased from Sigma-Aldrich or Fisher Scientific unless otherwise stated. Buffers and stock solutions were stored at room temperature unless otherwise specified.

2.1.1 Buffers

Phosphate Buffered Saline (PBS):

1 tablet of Dulbecco's PBS (Thermo Scientific) was dissolved per 100mL ddH_2O to make a solution of 1x PBS. To make a 10x stock solution, 10 tablets of Dulbecco's PBS were dissolved per 100mL ddH_2O. Solutions were autoclaved to sterilise.

Phosphate Buffered Saline + 0.05% Tween 20 (PBST):

5 Dulbecco's PBS tablets (Thermo Scientific) were dissolved in 500mL ddH_2O, and 250µL of Tween-20 was added to a final concentration of 0.05% (v/v), stirred thoroughly to dissolve.

10mM Sodium Citrate Heat Induced Epitope Retrieval (HIER) Buffer (pH 6.0)

2.94g tri-sodium citrate (dihydrate) was dissolved in 1L ddH_2O and stirred to dissolve. pH was adjusted to 6.0, 500µL of Tween-20 was added and stirred thoroughly.

Tris/EDTA HIER Buffer (pH 9.0)

1.21g Tris base and 0.37g di-sodium EDTA was dissolved in 1L ddH_2O and stirred to dissolve. pH was adjusted to 9.0, then add 500µL of Tween-20 and mixed.

Immunohistochemistry (IHC) /Immunofluorescence (IF) Blocking Buffer:

1mL fetal bovine serum (Sigma-Aldrich) to 10% (v/v), 0.5% (w/v) bovine serum albumin, 1mL 10x PBS, 8mL ddH_2O. Used immediately or stored at 4ºC until use.

10% Neutral Buffered Formalin:

90mL ddH_2O, 1 PBS tablet. Stir to dissolve, then add 10mL Formalin (37-40% formaldehyde solution stabilised with methanol) (Sigma-Aldrich).

Mowiol 4-88 Mountant:

2.4g Mowiol 4-88 (Calbiochem) was added to 6g glycerol and stirred to mix. 6mL ddH_2O was added and left to stir for 2 hours. 12mL 200mM Tris (pH 8.5) was added and the solution then went through repeated cycles of heating to 50ºC for 10 minutes and stirring for 10 minutes, until all the Mowiol had dissolved. The solution was clarified by centrifugation at 5000*g* for 15 minutes. 2.5% (v/v) Dabco 33-LV (Sigma-Aldrich) was added to the supernatant. Aliquoted and stored at -20ºC in the dark until use.

2.1.2 Antibodies

Table 2.1: Primary antibodies used in this thesis

Antibody	Host Species	Source (product code)	Application/ Dilution
ZSCAN4	Mouse	Novus Biologicals (H00201516-B01P)	IHC-P (IHC, paraffin sections) – 1:200 (Citrate HIER)
Nestin	Mouse	Millipore (MAB5326)	IHC-P – 1:200 (Citrate HIER)
BMI1	Mouse	Millipore (05-637)	IHC-P – 1:50 (Tris/EDTA)
ALDH7A1	Rabbit	Abcam (ab80187)	IHC-P – 1:200 (Tris/EDTA HIER)
AR	Rabbit	Santa Cruz Biotechnology (sc-816)	IHC-P – 1:1000 (Tris/EDTA HIER)
MUC1-C	Mouse	KeraFAST (EDF002)	IHC-P – 1:1000 (Citrate HIER)
Cytokeratin-5	Rabbit	Abcam (ab24647)	IHC-P – 1:1000 (Citrate HIER)
Syndecan-1	Mouse	Novus Biologicals (NB100-64980)	IHC-P – 1:200 (Citrate HIER)
Syndecan-1	Rabbit	Santa Cruz Biotechnology (sc25590)	IHC-P – 1:50 (Citrate HIER) IF-P (Immunofluorescence, paraffin sections) –

			1:50 (Citrate HIER)
SLC31A1	Rabbit	Novus Biologicals (NB100-402)	IHC-P – 1:250 (Tris/EDTA HIER)
RS1	Rabbit	Sigma-Aldrich (SAB1100315)	IHC-P – 1:500 (Citrate HIER)
p63	Mouse	Abcam (ab735)	IHC-P – 1:50 (Tris/EDTA HIER)
PSA	Rabbit	DAKO (A0562)	IF-P – 1:50 (Citrate HIER)
CD31	Mouse	DAKO (M0823)	IF-P – 1:30 (Citrate HIER)
S100	Rabbit	DAKO (Z0311)	IF-P – 1:2000 (Citrate HIER)
N-Cadherin	Rabbit	Abcam (ab18203)	IF-P – 1:100 (Citrate HIER)
E-Cadherin	Mouse	Zymed (33-4000)	IF-P – 1:25 (Citrate HIER)
Vimentin	Mouse	Santa Cruz Biotechnology (sc-620)	IF-P – 1:50 (Citrate HIER)
Pan Cytokeratin	Mouse	Sigma-Aldrich (C2931)	IF-P – 1:100 (Citrate HIER)
CD45	Rabbit	Santa Cruz Biotechnology (sc-25590)	IF-P – 1:50 (Citrate HIER)
CD19	Rabbit	Abcam (ab134114)	IF-P – 1:250 (Citrate HIER)
CD27	Rabbit	Abcam (ab131254)	IF-P – 1:100 (Citrate HIER)

Table 2.2: Secondary antibodies and dyes used in this thesis

Antibody/Reagent	Host Species	Source	Application/Dilution
Anti-Mouse IgG (H+L), Alexa Fluor 488 conjugate	Goat	Invitrogen (A-11001)	IF-P – 1:200
Anti-Rabbit IgG (H+L), Alexa Fluor 568 conjugate	Goat	Invitrogen (A-11008)	IF-P – 1:200
EnVision+ Anti-Mouse immunoglobulins, Horseradish Peroxidase (HRP)-labelled polymer conjugate	Goat	DAKO (K400611-2)	IHC – Ready to use
EnVision+ Anti-Rabbit immunoglobulins, HRP conjugate	Goat	DAKO (K401011-2)	IHC – Ready to use
4',6-diamidino-2-phenylindole (DAPI)	N/A (Nuclear probe)	Sigma-Aldrich (D9542)	Stock conc. – 1mg/mL Working conc. – 1µg/mL

2.2 Methods

All incubations and reactions were performed at room temperature unless otherwise stated.

2.2.1 Differential Gene Expression Analysis

2.2.1.1 Base methodology

Differential gene expression analysis on archived RNA microarray datasets was conducted using R version 3.2.0. Expression set data was acquired through the Gene Expression Omnibus database directly in R using the package *GEOquery* (ver. 2.34.0), and samples were filtered into categories based on the phenotypic data (e.g. Recurrence Status) or on expression levels based on a single probeset in the array (see Extremes of Expression Analysis). Differentially expressed genes between the groups were identified by fitting a linear model to the data using the *limma* package (ver 3.24.10) and the *eBayes* function within this package was used to generate empirical Bayes test statistics for each probe within the array. The function *topTable* was used to report the results, sorting by adjusted P-value (False Discovery Rate (FDR)) of the test for each probe to produce a list of top differentially expressed genes. The Benjamini and Hochberg (1995) FDR correction was used to account for multiple testing. An FDR<0.05 was considered statistically significant differential expression. Heat maps were generated for significant probes using the *heatmap.2* function from the *gplots* package (ver. 2.17.0), hierarchical clustering was performed and samples were labelled using the initial filtering criteria.

2.2.1.2 Extremes of Expression Analysis

To analyse the changes in gene expression that occur at the extremes of gene expression of candidate genes, differential gene expression analysis was performed as above, using the 10^{th} and 90^{th} percentiles of normalised signal of a probe to dichotomise the samples into ‘low’ and ‘high’ expression groupings and exclude data lying in the intermediate ranges from the analysis. A list of statistically significant genes meeting the FDR<0.05 criterion were passed into downstream gene ontology analysis and exported in a tabular CSV format, and a heat map was constructed to visualise differential expression patterns as described above. R code for this analysis technique is listed in Appendix 1.

2.2.1.3 Gene Ontology (GO) Analysis

GO analysis for lists of differentially expressed genes from Extremes of Expression Analysis was performed in R version 3.2.0 using the TopGO package (ver. 2.20.0). A set of custom functions listed top overrepresented GO terms for up-regulated and down-regulated genes in patient subsets highly expressing a candidate gene. A list of genes whose probes were present on the array platform were extracted from the microarray expression set and duplicates and probes not mapping to genic regions were discarded. The resulting gene list was known as the 'gene universe', and was used as the background set for comparison with differentially expressed gene lists. Current GO annotations for each gene in the gene universe were extracted from the ensembl database using the biomaRt package (ver 2.24.0). Differentially expressed gene lists were divided into up-regulated and down-regulated genes by separating them on the basis of log fold-change >0 and <0. Fisher's Exact tests were then used to compare the GO terms mapped to each gene list with GO terms in the gene universe, and GO terms with significant enrichment ($p<0.05$) in the differentially expressed gene list were output in a tabular CSV format. R code for this analysis technique is listed in Appendix 1.

2.2.2 Immunohistochemistry (IHC) on Formalin-Fixed Paraffin-Embedded (FFPE) Tissue

2.2.2.1 Coating of glass microscope slides with silane

To ensure robust adhesion of tissue sections to glass slides during antigen retrieval for immunohistochemistry and immunofluorescence, slides were coated with (3-aminopropyl)triethoxysilane (APTS). Slides were cleaned in absolute acetone for 2 minutes before allowing to dry in a fume hood. They were then immersed in a solution of 2% (v/v) APTS in absolute acetone for 2 minutes, washed in two changes of absolute acetone for 2 minutes each and left to air-dry again. Silane-coated slides were kept in the original slide boxes to keep them free of dust and debris before use.

2.2.2.2 Sectioning and staining method

FFPE tissue blocks were chilled to 4ºC before use, and 5µm thick sections were cut using a Leica RM2155 microtome or a Jung Biocut microtome fitted with disposable low-profile microtome blades (Fisher Scientific). Sections were floated out onto a 45ºC water bath, allowed to flatten, and then picked up onto silane-coated glass microscope

slides with a frosted edge. Sections were allowed to air-dry on a hot-plate set to 37ºC overnight before storage at room temperature or use in downstream staining protocols.

Prior to IHC staining, slides were deparaffinised in two changes of Histo-Clear for 2 minutes each, and Histo-Clear was removed from the sections using two changes of absolute ethanol for 1 minute each. Sections were rehydrated in a graded ethanol series by immersing for 1 minute in each solution: 95% (v/v) ethanol in ddH_2O; 90% (v/v) ethanol in ddH_2O; 70% ethanol in ddH_2O; 50% ethanol in ddH_2O. A final step of 2 minutes in ddH_2O was added to complete hydration.

To begin IHC staining, sections were washed twice in PBST. Nuclei were permeablised by incubation with 0.5% (v/v) Triton X-100 in PBS for 30 minutes, washed twice in PBST, and subjected to heat-induced epitope retrieval for 30 minutes at 90ºC in sodium citrate buffer (pH 6.0) or tris/EDTA buffer (pH 9.0) in a water bath, as optimised for the antigen to be detected. Slides were allowed to cool in the buffer bath at room temperature for 20 minutes and washed twice with PBST. Areas of tissue were circled with an ImmEdge hydrophobic barrier pen (Vector Laboratories) to ensure adequate coverage of reagents, and endogenous peroxidase activity was quenched with DAKO EnVision+ Peroxidase Block for 10 minutes. After washing twice with PBST, non-specific antibody binding was blocked with IHC blocking buffer for 30 minutes. Excess blocking buffer was drained off with a tissue and primary antibodies were added to the sections at the appropriate dilution (Table 2.1) in DAKO Antibody Diluent. Primary antibodies were incubated with the sections overnight at 4ºC in a humidified chamber.

The next day, sections were washed three times with PBST, and ready-to-use goat anti-mouse or goat anti-rabbit polymer-HRP-conjugated secondary antibodies were applied from the DAKO EnVision+ system (Table 2.2) for 30 minutes with gentle agitation on a rocking platform. Sections were washed three times, and signal was developed with the HRP substrate 3,3'-Diaminobenzidine (DAB) chromogen for 5 minutes, following the instructions according to the DAKO EnVision+ kit. Sections were washed three times with ddH_2O to terminate the reaction. Nuclei were counterstained with Gill's Haematoxylin (Vector Laboratories) for 1 minute, sections were rinsed under running tap water for 3 minutes, then differentiated by three dips in 1% (v/v) HCl in 70% (v/v) ethanol and 'blueing' in 0.25% (v/v) NH_3 in 70% (v/v) ethanol. Sections were dehydrated by going through the graded ethanol series in reverse order to finish with two changes of absolute ethanol. Tissues were cleared in two changes of Histo-Clear for

2 minutes each before mounting in DPX (Sigma-Aldrich) under a 22x50mm glass coverslip. Mounted slides were allowed to dry overnight in a fume hood before visualising under a Nikon Eclipse E800 widefield microscope using brightfield illumination. Images were captured on an attached Nikon Digital Sight DS-U1 CCD camera using NIS-F or NIS-D software.

2.2.3 Scoring of Immunohistochemical Stains

To quantify IHC staining patterns, a maximum of 15 random fields under a 20x objective were acquired from each tissue sample under each stain. 5 images were selected at random for quantification, although images had to contain epithelium to be considered for analysis. In some cases, biopsied tumour tissue was scant and in these cases, a minimum of 2 epithelium-containing images were required, or samples were excluded from the assessment. Stains were quantified according to the localisation of their staining patterns: BMI1, ZSCAN4, AR, ALDH7A1 and RS1 displayed epithelial nuclear staining patterns and were quantified by the Histo-scoring method (H-Score). The following cytoplasmic stains were scored using a cytoplasmic scoring method: Nestin, ALDH7A1, RS1, SDC1.

For H-Score assessment, all epithelial nuclei were counted and sorted into four different arbitrary categories based on observed DAB staining intensity: 0 – no staining; 1 – weak staining; 2 – intermediate staining; 3 – strong staining. The H-score $H$was calculated by applying the following formula, where i is the intensity category and P_i is the percentage of nuclei in an image in intensity category i as a proportion of total nuclei. The H-score can take any value from 0 to 300.

$$H = \sum_{i=1}^{3} iP_i$$

For quantification of cytoplasmic staining, the intensity and proportion of cells stained was estimated subjectively by placing the images into categories. Intensity categories were: 0 – no staining; 1 – weak staining; 2 – intermediate staining; 3 – strong staining. Proportion categories were: 0 – no staining; 1 - <25% epithelium stained; 2 – 25-50% epithelium stained; 3 – 50-75% epithelium stained; 4 – 75-100% epithelium stained. For SDC1 staining, where the proportion of cells stained varied, the Cytoplasmic Score was derived from the sum of intensity and proportion scores for a range of 0, 2-7. Where

stains varied only in intensity and 100% of epithelial cells were consistently stained, Cytoplasmic Score was derived from intensity score only for a range of 0-3. The following cytoplasmic stains were scored using the cytoplasmic intensity score: Nestin, ALDH7A1, RS1.

The burden of prostate cancer SDC1+ stromal positive (PCSP) cells was quantified in SDC1-stained prostate tissue samples by counting the number of PCSP cells per field, and measuring their area and shape parameters using ImageJ (NIH), calibrating the software to measure actual size by scale.

The means of every score from each set of five images was calculated and assigned to the sample as its overall score for each stain, regardless of score type. Differences in scores between relapsing and non-relapsing groups and benign and malignant tissue were statistically tested using Student's t-tests in R version 3.2.0.

2.2.4 Haematoxylin and Eosin (H&E) Staining of FFPE Tissue

FFPE tissue blocks were chilled to 4ºC and sections were cut at 5µm thickness, floated out onto silane-coated slides and air-dried as before. Prior to staining, sections were deparaffinised and rehydrated as described previously. Nuclei were stained with Gill's Haematoxylin (Vector Laboratories) for 2 minutes, rinsed with running tap water for 3 minutes, differentiated in 1% (v/v) HCl in 70% (v/v) ethanol for 30 seconds and 'blued' in 0.25% (v/v) NH_3 in 70% (v/v) ethanol for 1 minute. Sections were washed in 70% ethanol for 1 minute before counterstaining in 0.5% (w/v) Eosin Y (Raymond A Lamb) in ddH_2O for 5 minutes. This was followed by immersion in 95% (v/v) ethanol for 15 seconds, two changes of absolute ethanol for a minute each, and clearing in two changes of Histo-Clear for 2 minutes each. Excess Histo-Clear was blotted off of the slides and mounting was then achieved under a 22x50mm coverslip using DPX. Mounted slides were allowed to dry overnight in a fume hood before imaging using a Nikon Eclipse E800 widefield microscope under brightfield illumination, as described previously.

2.2.5 Immunofluorescence (IF) on FFPE Tissue

FFPE tissue blocks were chilled to 4ºC and sections were cut at 5µm thickness, floated out onto silane-coated slides and air-dried as before. Prior to staining, sections were deparaffinised and rehydrated as described previously. Washing, permeablisation and antigen retrieval was all performed as described previously. Sections were ringed with an ImmEdge hydrophobic barrier pen (Vector Laboratories) and blocked for 30 minutes

using IHC blocking buffer as in IHC, before incubation with primary antibodies (Table 2.1) diluted in DAKO Antibody Diluent (DAKO) overnight at 4ºC in a humidified chamber. For double immunolabelling experiments, two antibodies recognising different antigens and raised in different host species were mixed together at their respective dilutions for simultaneous labelling.

The next day, sections were washed three times in PBST and fluorescent labelling was performed in the dark for 30 minutes at room temperature with gentle agitation. Alexa Fluor 488-labelled goat anti-mouse IgG and Alexa Fluor 568-labelled goat anti-rabbit IgG antibodies were used, and both fluorescently labelled antibodies (Table 2.2) were mixed together at 1:200 dilutions in DAKO Antibody Diluent (DAKO) for simultaneous labelling. After incubation, sections were washed three times in PBST and then a modified protocol from Viegas et al. (2007) was used to quench endogenous tissue autofluorescence: samples were incubated with a filtered solution of 0.1% (w/v) Sudan Black B in 70% (v/v) ethanol for 10 minutes to quench endogenous tissue autofluorescence. Sections were rinsed briefly with PBST to remove any stain debris, then washed three times for 5 minutes each before mounting in Mowiol with 600nM DAPI (Sigma-Aldrich) for nuclear counterstaining. Samples were allowed to dry overnight in the dark at room temperature, or for at least two hours at 37ºC, before imaging on an LSM510 META laser scanning confocal microscope. Imaging was conducted using 405nm, 488nm and 543nm laser lines for excitation of fluorescent dyes and fields of view were collected under a 40x objective. Samples were stored at 4ºC in the dark and imaged within several days of staining for best results.

2.2.6 Tissue microarray (TMA) preparation and construction

2.2.6.1 Specimen acquisition and preparation

All tissues except prostate were FFPE archival specimens collected by Dr Asha Recino in September 2007. Mouse prostate tissues were harvested from adult CD-1 mice at 10 weeks of age after killing by a schedule 1 method. Whole urogenital sinuses were dissected from male mice and the prostate was isolated using the method detailed in Lukacs *et al.* (2010a), where ureters, fat, seminal vesicles, bladder and ureters are sequentially dissected away with forceps. Tissues were fixed for 24 hours in neutral buffered formalin, then processed using standard histological procedure. Samples were dehydrated in a graded ethanol series ranging from 70%-100% (v/v) concentration for 2 hours in each station; cleared in two changes of HistoClear for 2 hours each; and

infiltrated with three changes of HistoPlast paraffin wax (Thermo Scientific) at 65ºC for 2 hours each. Prostates were embedded whole, ventral side down, into molten HistoPlast wax and allowed to set.

2.2.6.2 TMA construction

Tissue microarrays were constructed manually using 1mm dermal biopsy punches (IHCWorld) inlaid into a premade recipient block (UNITMA) of 10x12 1mm diameter holes. A single tissue core of 1mm in diameter was taken from each specimen, guided by identification of representative regions on an H&E-stained whole section of the corresponding tissue block. Cores were inlaid into the tissue block, with blank remaining holes being filled with cores containing HistoPlast paraffin wax only. The TMA was processed according to manufacturer's instructions. A filled recipient block was heated to 70°C in an embedding mould on a hotplate until the array matrix was transparent, and molten HistoPlast paraffin wax and an embedding cassette was added as per the standard wax embedding protocol. The completed TMA contained 48 cores of mouse tissue from 13 different organs, and could be sectioned and stained using routine histological methods described in this thesis.

3 IDENTIFICATION OF POTENTIAL BIOMARKERS FOR RELAPSING PROSTATE CARCINOMA

3.1 Introduction

In order to identify new biomarkers for prostate carcinoma, a strategy for selecting potential biomarkers to test must be identified. Candidate markers must be selected and initially tested and characterised. This section will describe some of the recent progress in biomarkers for prostate cancer and review methods for identifying new biomarkers.

3.1.1 Recent progress in prostate cancer biomarkers

The National Institutes of Health have defined a biomarker as: "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmaceutical responses to a therapeutic intervention" (Bensalah *et al.*, 2007). Though there are many forms of biomarker, this thesis will consider a biomarker as a feature – commonly molecular features such as DNA, RNA or protein – that indicates a disease state. This section will summarise the different types of biomarkers in prostate cancer, and review some of the recent developments in prostate cancer biomarkers.

3.1.1.1 An overview of biomarkers

There are several main classes of biomarkers in prostate cancer. Diagnostic biomarkers indicate the presence or absence of prostate cancer in a patient, such as the previously mentioned serum PSA test (see Chapter 1.2.2). Prognostic biomarkers provide information about the predicted disease course of a patient with prostate cancer and help to stratify patients into poor outcome and favourable outcome categories (Bensalah *et al.*, 2007). The current best examples of clinical use of prognostic prostate cancer biomarkers are Gleason grade, tumour staging and PSA velocity, all covered in Chapter 1.2.2. Therapeutic target biomarkers identify the best courses of therapy for a patient (Bensalah *et al.*, 2007), and currently there are no such biomarkers in clinical use for prostate cancer. However, there are recent progresses in this area, with ligand-independent splice variants of AR such as AR-V7 being functionally implicated in castration resistance in patients undergoing hormone therapy (Hu *et al.*, 2009). As the aim of this thesis is to identify a prognostic biomarker, this section will focus on prognostic biomarkers. Molecular biomarkers are commonly assayed at several levels in patient samples: serum, urine, or tissue. Each carry advantages and disadvantages, as reviewed below.

3.1.1.2 Serum biomarkers

Serum biomarkers can give information as to the biological processes active within a tumour, without the need for a tissue biopsy. For example, shedding of the ectodomain of heparan sulphate proteoglycan Syndecan-1 (SDC1) is mediated by matrix metalloproteinases, and high levels of the shed SDC1 in patient serum are associated with metastasis and high Gleason grade tumours, and independently predict both shorter overall and disease-specific survival (Szarvas *et al.*, 2016a). Serum biomarkers can also

allow for the monitoring of normal organ function in the context of prostate carcinoma, which can provide useful information that tissue biopsies cannot. Patients that have been treated with radical prostatectomy can be separated into prognostic groups based on serum levels of the enzyme butyrylcholinesterase, which is synthesised in the liver and decreases in patients with liver damage or inflammatory conditions (Koie *et al.*, 2016). Patients with low levels of the enzyme prior to surgery have a 55% biochemical recurrence-free survival rate, whereas patients with high levels have a 77.7% rate. It suggests that patients may develop conditions secondary to prostate carcinoma that cause such a response, although the exact causes of this elevation are not known (Koie *et al.*, 2016).

Another advantage of serum biomarkers is that they can be analysed in the absence of tumour material needed for a biopsy, such as in a screening programme. Elevated plasma levels of insulin-like growth factor IGF-1 and its binding protein IGFBP-3 were found to be associated with an increased risk of developing advanced stage prostate carcinoma, even though the blood samples were taken on average 9 years before a diagnosis of carcinoma was reached (Chan *et al.*, 2002). Given the natural history of prostate carcinoma detailed in Chapter 1.2, it is possible that a biomarker could function as an early predictor of carcinoma in later life. Unfortunately, IGF-1 and IGFBP-3 were inadequate diagnostic tools, underperforming PSA in the diagnosis of carcinoma, which limits their utility. However, the study highlights the power of serum biomarkers, where samples can be taken even before a tumour is identified and thus they make ideal screening tools. Since systemic levels of the analyte are measured, there are fewer sources of error introduced by sampling: tissue biopsies normally sample only a small fraction of the prostate, which can lead to misrepresentation of the overall tumour mass (King and Long, 2000). There are no serum prognostic biomarkers currently in use, and it is a future area of interest and potential.

3.1.1.3 Urinary biomarkers

Urine provides another minimally invasive source of biomarkers for prostate cancer. However, this field has yet to yield significant amounts of potential prognostic biomarkers. PCA3, a prostate cancer-related antigen, is established for the detection of prostate cancer using urine (Sokoll *et al.*, 2008; Merola *et al.*, 2015). Following digital rectal examination, both PCA3 mRNA levels are quantified in the urine and normalised to PSA mRNA, presumably from liberated cells of the prostate epithelium post-

examination. The method is suitably sensitive and specific for the detection of prostate cancer (Sokoll *et al.*, 2008; Merola *et al.*, 2015) and is currently considered for diagnostic use. However, it is unclear whether PCA3 may also be a prognostic biomarker. Urinary PCA3 levels may positively correlate with Gleason grade of primary tumours (Merola *et al.*, 2015; Wei *et al.*, 2015) and with the presence of advanced stage tumours (Wei *et al.*, 2015), although other studies have found no association with Gleason grade (Augustin *et al.*, 2013) and there is still controversy regarding the association between urinary PCA3 mRNA levels and measures of tumour aggressiveness such as Gleason grade or stage (Augustin *et al.*, 2013; Leyten *et al.*, 2014).

Recent studies have found that a similar detection of TMPRSS2-ERG fusion mRNA in urine can be combined with PCA3 mRNA tests in urine to provide a more specific cancer diagnostic ability than PSA assessment alone (Tomlins *et al.*, 2016; Leyten *et al.*, 2014), especially when detecting high-grade prostate cancer (Tomlins *et al.*, 2016). TMPRSS2-ERG mRNA levels positively correlate with Gleason grade and tumour stage in the urine, although PCA3 mRNA may not share this attribute (Leyten *et al.*, 2014). Because of the minimal invasiveness of urine collection, it is possible to adapt these assays for use in a primary care setting: a sensitive test for TMPRSS2-ERG fusion transcripts in urine has been developed using an isothermal transcription-mediated amplification method (Koo *et al.*, 2016), which requires no specialist equipment and has an easily interpretable binary output. This offers many advantages, including potentially speeding up decision-making processes in primary care, and would be invaluable if applied to a suitably prognostic biomarker. However, the urinary biomarker field has yet to yield a significant amount of potential urinary prognostic biomarkers. It remains to be determined if TMPRSS2-ERG, PCA3 or any other marker can fulfil such roles, especially as long-term follow-up information on recurrence and survival is absent in all studies so far.

3.1.1.4 Tissue biomarkers

Despite the benefits of fluid biopsies such as urine and serum, there are several advantages to tissue biomarkers. IHC techniques for tissue biomarkers are already in use in pathology laboratories for diagnostic applications, meaning that inclusion of a new antibody for staining is simple, existing equipment may be used, and existing core needle biopsies may be used, which reduces the need for additional procedures for

sample collection. Additionally, they provide protein localisation data that may help to inform a diagnostic or prognostic prediction. For example, the fibroblast growth factor homologous factor FGF13 is elevated at the protein level in prostate cancer compared to high-grade PIN and normal tissue (Yu *et al.*, 2016). Whereas cytoplasmic FGF13 IHC staining is associated with shorter recurrence-free survival, low cytoplasmic staining and high nuclear staining is associated with a much better prognosis (Yu *et al.*, 2016). IHC therefore has the potential to provide much more information about patient outcome than measurement of a single level in urine or serum, because both protein levels and localisation can be observed.

Another role of tissue biomarkers is in the molecular subtyping of tumours. A good example of this is in breast cancer, where the IHC detection of estrogen receptor, progesterone receptor and HER2 allows for the classification of patients into molecular subtypes and tailored treatment (Dent *et al.*, 2007). For example, patients with no expression of the three markers are in the poor prognostic category known as 'triple-negative' breast cancer and will not benefit from tamoxifen or Herceptin therapies, whereas patients with HER2-expressing tumours will benefit from Herceptin, and hormone receptor-positive patients will benefit from tamoxifen (Dent *et al.*, 2007).

There have been recent advances in prostate cancer subtyping with the finding of ETS-ERG fusions and SPOP mutations (Abeshouse *et al.*, 2015), as well as with SPINK1 overexpression (Tomlins *et al.*, 2015) that might support such a panel in prostate cancer. In patients with initially diagnosed bone metastatic prostate carcinoma, expression of the serine protease inhibitor SPINK1 at the protein level was detected in 13.5% of patients (Pan *et al.*, 2016) and indicated a 1.7-fold greater risk of progression to castration-resistant prostate cancer. Similarly, in patients treated with primary hormone therapy, SPINK1 protein expression is found in 10% of tumours and is associated with disease recurrence (Leinonen *et al.*, 2010). Conversely, other studies of SPINK1 expression in patients undergoing prostatectomy found no association between SPINK1 and recurrence or Gleason grade (Grupp *et al.*, 2013; Flavin *et al.*, 2014), which might show evidence of treatment-specific differences in SPINK1 association. Interestingly, SPINK1 expression seems to be mutually exclusive with ERG expression, suggesting a separate subtype of prostate tumour (Pan *et al.*, 2016; Grupp *et al.*, 2013; Bismar *et al.*, 2012). SPINK1 overexpression may also be associated with PTEN-deletion in ERG-negative tumours (Bismar *et al.*, 2012), although this is controversial (Flavin *et al.*, 2014). Analysis of microarray data on large cohorts confirms that ERG+, ETS+ and

SPINK1+ are distinct molecular subtypes of prostate carcinoma, along with a fourth triple-negative classifier. Such molecular classification of prostate tumours apparently has very limited prognostic utility in prostate cancer (Tomlins *et al.*, 2015) compared with breast cancer. However, as this data is limited to radical prostatectomy patients, it is still possible that these classifications may provide additional information about the best course of treatment. Thus, a panel of biomarkers such as ETS/ERG, SPINK1 and PTEN antibodies for IHC could still be of clinical use.

Taken together, there are considerable advantages to the use of serum and urinary biomarkers in the diagnosis and prognostication of prostate cancer. However, the additional protein localisation information provided by IHC and the ease with which a new biomarker can be implemented in pathology laboratories makes tissue biomarkers an indispensable tool for prostate pathology.

3.1.2 Identification of new biomarkers

The first step in this project was to identify biomarkers to test. There are multiple methods of identifying biomarkers, each with individual strengths and weaknesses. Some common approaches are reviewed below.

3.1.2.1 Transcriptomics

An increasingly popular approach for identifying new biomarkers in large clinical cohorts is to investigate changes in the transcriptome of patient tumours, usually by cDNA microarray or high-throughput RNA sequencing. In many cases these biomarkers take the form of gene signatures, where the expression levels of multiple genes in a microarray are used to predict disease state. The first example of this was demonstrated in 1999 when a 50-gene signature obtained from microarray data was used to accurately classify patients as having Acute Myeloid Leukaemia or Acute Lymphoblastic Leukaemia, independent of other biological knowledge (Golub *et al.*, 1999). This approach has experienced some success in predicting recurrence (Sun and Goodison, 2009) and in predicting outcome in multiple tumour types (Glinsky *et al.*, 2005), including predicting therapy resistance in prostate cancer (Rajan *et al.*, 2014) as well as in differentiating disease subtypes in prostate cancer (Ben-Porath *et al.*, 2008). However, the approach relies upon multi-gene models and often requires higher throughput methods with which to evaluate patient samples. Integrative computational approaches which make use of microarray data combined with *a priori* biological

knowledge, pathway annotations and clinical data are increasingly being adopted, but require extensive validation and new modelling methods due to high data dimensionality compared to sample size (Iuliano *et al.*, 2016).

Removing single genes from multi-gene models in order to use them independently removes them from the context of their selection, and it is possible that genes used singly will not have the utility that the multi-gene model possessed. Due to over-identification, it is important that these biomarkers have independent experimental validation. However, these studies supply an ever-growing source of candidate markers for investigations to validate or to characterise at other levels, such as on the genomic or protein levels.

3.1.2.2 Proteomics

Protein-based assay techniques such as IHC and ELISA are important tools in diagnostic pathology. Biomarkers at the protein level have been identified in screens using proteomics on tumour tissue or urine samples and compared to benign counterparts, although this is still a comparatively underused approach compared to transcriptomics. For example, a Stable Amino Acid Labelling in Cell Culture (SILAC) and mass spectrometry technique was used to identify the biomarker proneuropeptide-Y as a marker of progression and poor outcome in primary prostate tumour tissue, and was validated in patient tissue microarrays using immunohistochemistry (Iglesias-Gato *et al.*, 2016). In contrast, a proteomics study conducted in prostate fluid and urine failed to identify a single biomarker that could distinguish patient subgroups (Kim *et al.*, 2016), suggesting that while the approach is effective in identifying biomarker panels, a single protein might not be sufficient for such a task. Another study investigated the proteome of exosomes in patient serum, and found an enrichment of known cancer-associated proteins as well as novel ethnicity-specific proteins that could serve as useful ethnicity-specific cancer biomarkers (Turay *et al.*, 2016). These studies demonstrate that the approach has promise, especially with respect to the profiling of liquid biopsies, which could serve as a minimally invasive source of biomarkers. As with transcriptomics assays, they provide a wealth of potential biomarkers for further follow-up studies.

3.1.2.3 Literature Searching

Literature searching to identify potential new biomarkers can be as simple as a series of keyword searches and manual examination and curation of the literature. The advantage of such approaches is that no knowledge of text-mining and no benchmarking of search

methodology is required. The knowledge and experience of the curator(s) is used to critically evaluate the literature while it is being collected, allowing for effective (albeit subjective) collection of relevant literature material for decision-making. However, such approaches are time intensive: one extensive literature searching approach to identify pancreatic cancer biomarkers took approximately 7000 person hours to complete (Harsha *et al.*, 2009).

The additional caveat of manual search approaches is that, because the search is time-consuming, by the time a literature review is completed there will be yet more currently published biomarkers to add to the list for review. This leads to a continuous build-up of evidence, and the rate of evidence accumulation far surpasses the rate at which the evidence can be tested and evaluated. This abundance of biomarker information is largely due to published 'omics' datasets which routinely identify overexpressed proteins in a variety of pathological situations. However, it is arguable that this abundance of information is difficult to objectively hone down into just a small number of useful biomarkers which would go forward for further study.

To resolve this problem, automated computer-based approaches have been devised to extract key information from published abstracts and main texts, such as gene and protein names, evidence categories and statistics (Younesi *et al.*, 2012). The information is ranked and scored in a defined, objective manner using algorithms. The search terms, sensitivity and specificity of the searching approach can be tailored to meet the needs of the researcher. Regardless, the increasing wealth of biomarker information means that a centralised database for storing and interrogating biomarker evidence is becoming increasingly important (Harsha *et al.*, 2009).

Such automated approaches are less time-intensive whilst requiring more skill in informatics, but it is still unclear whether there are other defined benefits to quantitation of the literature over a manual searching approach. Automated search approaches require manual verification and may eliminate useful literature and biomarker leads that a trained individual could manually identify, but do so systematically and consistently. The human element of a manual search is both an advantage and a disadvantage, as this can make the literature search less systematic and potentially result in the incorporation of additional sources of bias. An additional flaw of both manual and automated literature searching approaches is that they are not suitable for identifying novel combinations of biomarkers in multi-biomarker panels, which may be more accurate

and informative in classifying disease state than any single biomarker (Kim *et al.*, 2016). However, all literature searching approaches have the advantage of testing biomarkers that have already been studied to some degree by other experimental approaches, meaning that, while not strictly novel, a thorough literature search should result in an increased chance of identifying a successful biomarker.

3.1.2.4 The contribution of tissue microarrays to biomarker validation

Having identified potential biomarkers, tissue microarrays provide a means to section, stain and analyse a larger number of samples of FFPE or frozen tissue specimens. The throughput capability is higher than a whole section study can accomplish, which is achieved by arranging representative cores of tissue in an array format prior to sectioning. This allows routine IHC or in situ hybridisation procedures to investigate selected biomarkers in a higher throughput setting with reductions in time, cost and precious sample material (Hassan *et al.*, 2008). Multiple replicate cores can be included to better represent tissue heterogeneity in each specimen block. This technique ensures consistent and uniform sample treatment while also benefiting from the increased statistical resolving power that comes from testing a larger patient cohort. For example, Ruiz *et al.* used two tissue microarray blocks consisting of over a thousand prostate tumour samples together to evaluate NRBP1 (2012), and the Melling *et al.* study on HOOK3 expression in prostate cancer used more than 12,400 samples in their tissue microarray (2015).

The main disadvantage of using TMA technology to validate biomarkers is that, like all IHC studies, they require an antibody that performs reliably when used for the application. When these antibodies are ultimately destined to be used in a clinical pathology setting, this may also be considered as an additional filtering characteristic of the approach. There is also the problem of tissue size: the cores used in a tissue microarray can be as small as 0.6mm in diameter, and there is concern as to whether this is representative of the heterogeneity of prostate tumours when compared to whole tissue sections, especially where staining patterns are focal or heterogeneous (Merseburger *et al.*, 2003). This can be overcome by including more than one replicate core per specimen in a tissue microarray, ensuring that more of the heterogeneity is encompassed for a more representative result (Mucci *et al.*, 2000). Such limitations are also counterbalanced by the increased density of information available when such large

numbers of patients are tested simultaneously (Varambally *et al.*, 2002; Melling *et al.*, 2015).

Tissue microarrays are often accompanied by patient clinical data and are designed for multiple purposes tailored to the screening approach. For example, a tissue microarray for disease progression or diagnostic biomarkers might contain a mixture of normal prostate tissue, benign mimics, high-grade PIN and prostate carcinoma to ensure that the tested antibodies can distinguish those tissue types by IHC. A microarray of this type was recently used to show that the RNA-binding protein hnRNPA2 is more highly expressed in high grade tumours (Stockley *et al.*, 2014), and the same approach has been taken to characterise NRBP1 (Ruiz *et al.*, 2012).

Similarly, a tissue microarray for a prognostic biomarker might contain tumour samples from patients with a wide range of clinical outcomes, and include at least five years of clinical follow-up data. When such clinical data are available, it is possible to test the relationship between biomarker expression status and recurrence (Melling *et al.*, 2015; Ruiz *et al.*, 2012; Leclerc *et al.*, 2016) and other patient metadata, such as ERG rearrangement status (Melling *et al.*, 2015). An additional advantage of tissue microarrays is that adjacent sections can be stained with other antibodies and staining patterns examined and correlated. For example, by staining an adjacent section with an anti-Ki-67 antibody, which is a marker of proliferation, it is possible to identify associations between biomarker expression and Ki-67 expression, suggestive of an effect on cell proliferation (Ruiz *et al.*, 2012; Melling *et al.*, 2015).

In summary, while there are technical challenges to be overcome with the technology, the tissue microarray has numerous benefits for high-throughput analysis of patient cohorts for biomarker validation, and provides a useful platform that could be utilised in the current study for the purpose of distinguishing recurrent from non-recurrent patient tumours.

3.1.3 Aims

This chapter will begin the project by addressing the following initial aims:

i. Identify potential new biomarkers that be investigated in subsequent chapters, using two different methodologies:
    a. A literature search to identify promising candidates that require further characterisation, and;
    b. A bioinformatic approach using publically-available microarray datasets to identify other potential biomarkers using an unbiased, computational method.
ii. Develop and use bioinformatic methods to gather more information about the selected biomarkers and further inform the decision making processes.
iii. Establish tissue microarray construction and staining protocols that can be used to construct a tissue microarray from precious prostate biopsy samples to be analysed in future chapters.
iv. Perform initial antibody optimisation for IHC on this tissue microarray platform.

3.2 Results

3.2.1 Identification of potential biomarkers through literature review

The first step in identifying new biomarkers was to select candidate markers to test. Because of the versatility and ease of implementation of tissue biomarkers, it was decided to search for tissue biomarkers for this study. In order to identify candidate markers for testing, two approaches were used: a literature searching approach, and a bioinformatics approach.

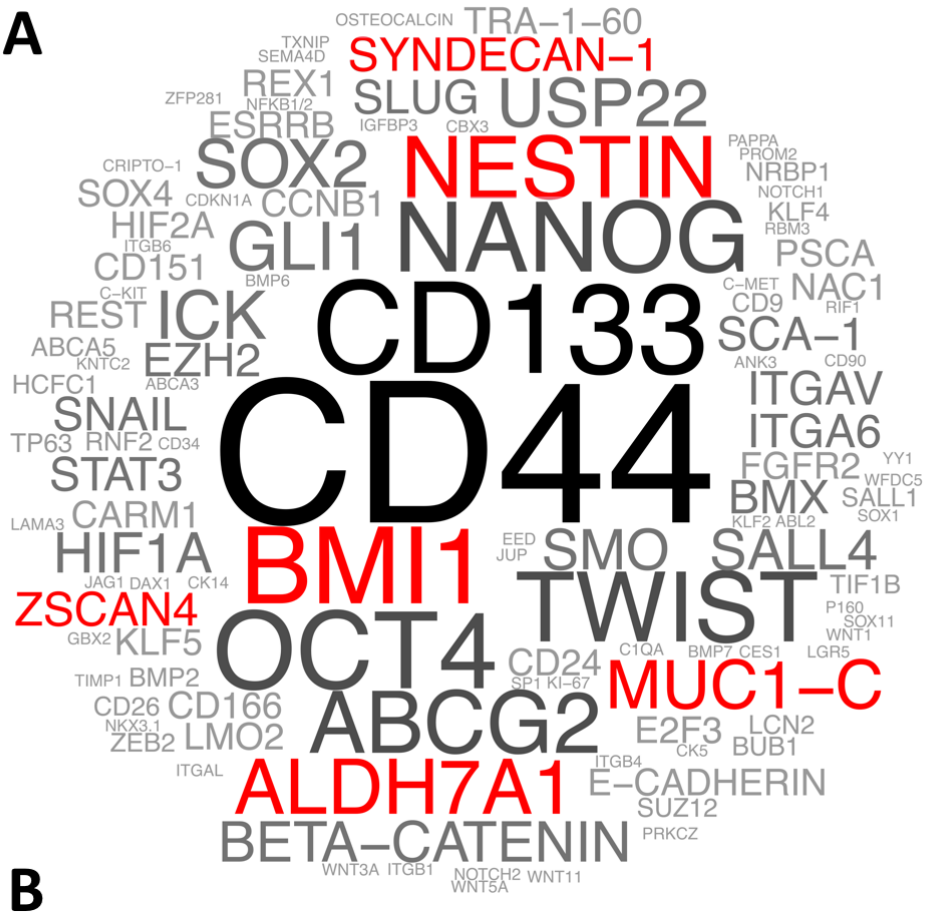
The first approach was to conduct a thorough literature review using Web of Science (www.webofknowledge.com. Date accessed: November 2012). A total of 116 potential biomarkers were identified (Figure 3.1A) based on their manually-curated relationship in the literature with five criteria. These were: prostate cancer; recurrence (in any cancer type); other non-prostate cancers; cancer stem cells; and tissue stem cells (Figure 3.1B). The search began with a generic search, using the terms "prostate cancer relapse" and "prostate cancer stem cell", branching out by including markers identified in this initial pass using a more specific search strategy: i.e. searches with the marker name and "prostate cancer", "cancer", "stem cell" and "cancer stem cell".

This search method was driven by the need to identify markers that may be important in tumour recurrence, but also to identify biomarkers that might be of wider relevance in the cancer field and gone unstudied in prostate cancer. For example, the Polycomb group protein BMI1 was first identified to be a haematopoietic stem cell-related protein (Alkema *et al.*, 1997), but further studies showed it to be involved in maintaining the intestinal stem cell niche (Sangiorgi and Capecchi, 2008) and that it was expressed in putative prostate stem cells as well as being important for malignant transformation (Lukacs *et al.*, 2010b). Similarly, tumours and stem cells often share expression characteristics which appear to correspond to levels of plasticity and differentiation state (Palmer *et al.*, 2012), and as stem-like expression signatures have been shown to be associated with poor outcomes in a variety of tumour types including prostate cancer (Ben-Porath *et al.*, 2008; Markert *et al.*, 2011), it was important to identify biomarkers that met these criteria.

I selected a range of candidate markers with different extents of characterisation and literature coverage, including well-studied candidates already flagged for biomarker use such as BMI1, and poorly-studied candidates whose expressions had not yet been reported in prostate cancer, such as ZSCAN4. This was done in order to spread risk

associated with a literature searching approach and increase the chances of success: allowing higher-risk, briefly-characterised candidates to be investigated for potentially novel findings (such as ZSCAN4 and ALDH7A1), while also having lower-risk, well-characterised candidates to study (such as BMI1 and SDC1). A long list of 116 candidate markers was identified (Figure 3.1A). From this, 12 were shortlisted for further consideration. The 12 shortlisted biomarkers, complete with curated evidence from the time of decision making in 2012-2013, are shown in Table 3.1. Of these, 6 candidates were chosen for experimental study, based on an assessment of the evidence presented for their utility as biomarkers: relationship to clinical parameters, including relapse status; and any interesting biological characteristics, such as a relationship to cancer stem cell behaviour, metastasis or hallmarks of cancer. These candidates were: ALDH7A1, BMI1, Nestin, MUC1-C, ZSCAN4 and SDC1. As before, candidates were chosen to represent differing degrees of coverage in the literature, ranging from candidates with prior links to clinical outcome (BMI1, SDC1, MUC1-C) and/or castration resistance (NES, MUC1-C) to candidates with no clinical associations in the literature but interesting biological characteristics (ALDH7A1, ZSCAN4). These markers are reviewed with current literature in Chapter 4.1.1.

A


OSTEOCALCIN
TRA–1–60
TXNIP
SEMA4D
SYNDECAN–1
REX1
SLUG
USP22
ZFP281
NFKB1/2
ESRRB
IGFBP3
CBX3
PAPPA
PROM2
CRIPTO–1
SOX2
NESTIN
NRBP1
SOX4
CDKN1A
CCNB1
NOTCH1
KLF4
HIF2A
RBM3
ITGB6
GLI1
NANOG
PSCA
CD151
BMP6
C–KIT
C–MET
NAC1
REST
ICK
CD133
CD9
RIF1
SCA–1
ABCA5
KNTC2
EZH2
ANK3
CD90
HCFC1
ABCA3
CD44
ITGAV
SNAIL
ITGA6
TP63
RNF2
CD34
FGFR2
YY1
STAT3
WFDC5
BMX
SALL1
LAMA3
CARM1
KLF2
ABL2
SOX1
HIF1A
BMI1
EED
JUP
SMO
SALL4
JAG1
DAX1
CK14
TWIST
TIF1B
ZSCAN4
P160
SOX11
GBX2
KLF5
OCT4
WNT1
C1QA
BMP7
CES1
LGR5
CD24
TIMP1
BMP2
SP1
KI–67
MUC1–C
CD26
CD166
ABCG2
NKX3.1
ZEB2
LMO2
E2F3
LCN2
CK5
BUB1
ITGAL
ITGB4
ALDH7A1
E–CADHERIN
SUZ12
BETA–CATENIN
PRKCZ
WNT3A
ITGB1
NOTCH2
WNT11
WNT5A

B

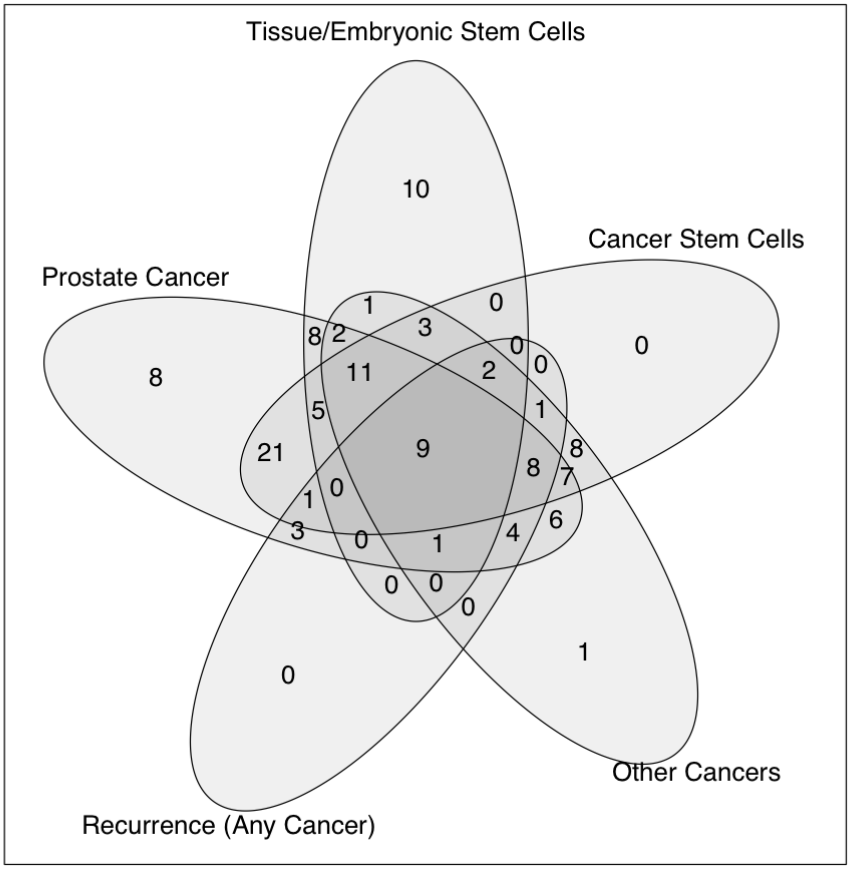

Tissue/Embryonic Stem Cells
Cancer Stem Cells
Prostate Cancer
Other Cancers
Recurrence (Any Cancer)
10
8
0
1
3
0
8
2
0
0
11
2
0
5
1
21
9
8
8
7
1
0
3
0
1
4
6
0
0
0
1
0

Figure 3.1: Graphical representation of literature searching results overall and by evidence category.

(A) Schematic overview showing all 116 potential biomarkers looked at in this study. The size of each protein name is proportional to the amount of literature identified that met the search criteria, with the biomarkers chosen for further study shown in red. Literature searching identified many types of markers, including stem cell markers such as NANOG and OCT4, canonical CSC markers CD44 and CD133, and EMT markers such as TWIST and SLUG. (B) Venn diagram showing classification of biomarkers by criteria they are associated with: prostate cancer; tissue/embryonic stem cells; cancer stem cells; recurrence (in any cancer); or relationships with other cancers.

Table 3.1: A sample of 12 shortlisted biomarkers for potential analysis.

A description of each marker is included underneath the marker name. Markers in **bold** are part of the six final candidates. Reasons for selection includes principle findings and key citations.

Marker name	*Reasons for selection*
BMI1 *Component of the polycomb complex PRC1. It monoubiquitylates H2AK119 in order to epigenetically regulate gene silencing during development and cancer.*	BMI1 expression is highly enriched in stem-like prostatospheres generated from multiple cell lines (Rybak *et al.*, 2011; Lukacs *et al.*, 2010b; Zhang *et al.*, 2012a), prostate cancer biopsies and invasive xenograft tumours and cell lines. BMI1 is overexpressed in prostate cancer biopsies and may be an independent prognostic indicator, predicting PSA recurrence (van Leenders *et al.*, 2007). Expression increases with increasing tumour stage in human prostate cancers (Siddique *et al.*, 2013a). It is also a biomarker for prostate cancer diagnosis in patient serum (Siddique *et al.*, 2013c).
MUC1 (MUC1-C) *Transmembrane Mucin. The C-terminal subunit has a function in many cell signalling processes involving cell survival, cell adhesion and cell proliferation. The C-terminal subunit can enter the nucleus and affect transcription. The N-terminal subunit serves as a mucous*	This protein is overexpressed in many cancers, particularly adenocarcinomas such as breast cancer (Engelmann *et al.*, 2008). Functionally, MUC1-C has many roles including in maintaining embryonic stem cell pluripotency (Hikita *et al.*, 2008). No studies suggesting a CSC link in prostate cancer, but expression of the C-terminal domain has been associated with relapse, high tumour grade and advanced tumour stage (Lapointe *et al.*, 2004). It represses androgen receptor expression and is

membrane protein through its highly glycosylated PTS repeat motifs.	involved in castration resistance and chronic inflammation (Rajabi *et al.*, 2012). Blocking oligomerisation of MUC1-C causes regression of xenograft prostate tumours (Joshi *et al.*, 2009).
Nestin *Intermediate filament protein. Important for brain and eye development, and involved in intermediate filament assembly and disassembly during neural differentiation, but lost upon terminal differentiation.*	Nestin was expressed in prostate cancer stem cells generated by immortalization of human prostate tumour cells (Gu *et al.*, 2007). Nestin is expressed in lethal metastatic CRPC (Kleeberger *et al.*, 2007a). Expression of nestin and Ki-67 together are associated with biochemical failure, recurrence and metastasis – but nestin alone is not associated with clinicopathological variables(Gravdal *et al.*, 2009). In another study, nestin is upregulated in metastatic CRPC when compared to non-castrated, metastatic prostate cancer (Pfeiffer and Schalken, 2010).
ALDH7A1 *Aldehyde Dehydrogenase. Involved in Lysine catabolism and resistance to oxidative stress.*	Aldehyde dehydrogenase activity identified a subset of cells with tumour-initiating capacity in prostate cancer cell lines (Hellsten *et al.*, 2011; van den Hoogen *et al.*, 2010; Qin *et al.*, 2012) when also sorted for other cancer stem cell markers such as CD44 and α2β1 integrin. Prostate CSCs isolated in this way displayed quiescence and resistance to chemotherapeutic agents and androgen deprivation (Qin *et al.*, 2012). ALDH7A1 is highly expressed in prostate cancers (van den Hoogen *et al.*, 2011) and is also in

	matched bone metastases. ALDH7A1 knockdown reduces bone metastasis in mouse xenografts. Therefore, ALDH activity is linked to cancer stem cell properties – and ALDH7A1 in particular is linked to metastasis.
ZSCAN4 *Transcription Factor, essential for embryonic stem cell pluripotency and also regulates telomere length in ES cells.*	ZSCAN4 is known to be necessary for pre-implantation development of mouse embryos and is expressed heterogeneously in mouse ES cells (Falco *et al.*, 2007). Exogenous expression of ZSCAN4 increases induction efficiency of induced pluripotent stem cells, and replaces Myc in their induction (Hirata *et al.*, 2012). This suggests it has a role in regulating stemness. Additionally, ZSCAN4 regulates telomere elongation in ES cells (Zalzman *et al.*, 2010),which is important for avoiding senescence in stem cells. However, ZSCAN4 expression has not been investigated thoroughly in adult tissues or cancer, with one paper (Ko *et al.*, 2013) suggesting that its expression may be amplified along with stem cell markers during inflammation of the pancreas.
SDC1 *Syndecan-family member of heparan sulphate proteoglycans on the cell surface, linking the cytoskeleton to the extracellular matrix.*	Altered Syndecan expression appears to be associated with a greater risk of PSA relapse in patients with clinically localized prostate cancer (Ledezma *et al.*, 2011; Chen *et al.*, 2004b). It is also associated with the expression of EMT markers (Contreras *et al.*, 2010).

ABCG2 *A member of the ATP-Binding Cassette (ABC) Transporter Family. Also known as BCRP, it may be involved in resistance to xenobiotics.*	Very well studied in many cancers, and often linked to cancer stem cells. It has mainly been characterized in prostate cancer cell lines via FACS. Thought to be responsible for the 'side population' phenotype whereby CSCs and other stem cells are resistant to chemotherapeutic agents and can efflux Hoechst 33342 dye (Zhou *et al.*, 2001). Stem-like prostatospheres generated from prostate cancer cell lines overexpressed ABCG2 along with CD44 and other cancer stem cell markers (Zhang *et al.*, 2012a) and were resistant to chemotherapeutic agents. In the murine prostate, adult stem-like prostate cells with high ALDH activity also highly express ABCG2 (Burger *et al.*, 2009).
USP22 *Part of the hSAGA coactivator complex, and has histone deubiquitinating activity.*	USP22 expression has been identified in a stem-like gene expression signature related to therapy failure in multiple cancers – including prostate cancer (Glinsky *et al.*, 2005). USP22 expression correlates with poor prognosis, especially with BMI-1 expression, in multiple cancers. High expression of USP22 in patients with is associated with poor clinical outcome in breast cancer (Zhang *et al.*, 2011) and gastric cancer (Yang *et al.*, 2011). In colorectal cancer, USP22 expression is elevated in metastases and its increased expression in is associated with poor response to therapy (Liu *et al.*, 2011b). As the prognostic value of USP22 is often

	improved when combined with expression of BMI1 (Yang *et al.*, 2011; Hu *et al.*, 2012), it is presumed to activate the BMI1 driven oncogenic pathways.
SALL4 *Probable zinc finger transcription factor.*	Related to the Hoechst side population, ABC transporter expression and drug resistance. May interact with Nanog in ES cells (Wang *et al.*, 2006). Expressed in early stages of breast cancer (Kobayashi *et al.*, 2011b) and lung cancer (Kobayashi *et al.*, 2011a) independent of tumour stage. Overexpression and knockdown of SALL4 in HEK293 cells leads to increased and decreased drug resistance and expression of ABC transporters (Jeong *et al.*, 2011). Overexpression of SALL4 in MCF7 breast cancer cells results in an expansion of the stem-like side population. In the prostate cancer cell line PC-3, stem-like cells expressing PDGF-D have upregulated SALL4 expression (Kong *et al.*, 2010), along with increased expression of markers of EMT.
E2F3 *E2F transcription factor. Important in cell-cycle control and tumour suppression.*	Related to prostate stem cells as it is upregulated in the mouse urogenital sinus, which is enriched in stem cells and forms the prostate during development (Blum *et al.*). Expression of E2F3 negative correlates with survival in prostate cancer patients, with highly expressing tumours having a poor prognosis independent of stage (Foster *et al.*) E2F3 has been identified as part of a panel of 5 genes that predict

	biochemical recurrence in prostate cancer patients (Barwick *et al.*, 2010).
ITGAV *αV Integrin. Interacts with the Extracellular Matrix and also mediates signal transduction. Receptor for Vitronectin and a variety of other ligands containing the peptide sequence RGD.*	αV Integrin is enriched in CD44+/CD24- breast cancer stem cells when compared to the non-CSC population (Murohashi *et al.*, 2010). αV Integrin is involved in the response of prostate cancer stem cells to vitronectin, which is the component of serum that causes CSC-like cells to differentiate in vitro (Hurt *et al.*, 2010). Conversely, culturing PC-3M cells with an αV integrin antagonist decreases the stem-like ALDH-high population and attenuates metastasis in another study (van der Horst *et al.*, 2011).
STAT3 *Transcription Factor. Involved in the response to interleukins and acute phase proteins. STAT3 is activated through the JAK-STAT pathway.*	Inhibition of STAT3 reduces CSC phenotypes in stem-like colon cancer cells (Hellsten *et al.*, 2011) and breast cancer xenografts (Dave *et al.*, 2012). STAT3 is upregulated in stem-like prostatosphere cells and in invasive prostate carcinoma (Mathews *et al.*, 2010). STAT3 inhibition reduced invasion in prostate cancer cell lines. However, STAT3 function with regard to the expression of stem cell related genes has not been investigated.

3.2.2 Identification of candidate biomarkers through bioinformatics

The second approach to identify candidate biomarkers was a bioinformatics approach. This provided an unbiased approach which complemented the literature searching approach. To do this, I used a cDNA microarray dataset (GEO Accession: GDS4109) published by Sun and Goodison (2009), which contains 79 primary prostate tumour samples divided into two categories: 39 were classified as recurrent, defined as biochemical recurrence with a rising PSA at three consecutive measurements, whilst 40 were non-recurrent in the first 5 years of follow-up. Samples were log2-transformed and differential gene expression analysis was conducted between these two groups of patients in order to identify genes that were significantly differentially expressed between recurrent and non-recurrent patient tumours. The hypothesis was that differentially expressed genes between these two groups could be candidate biomarkers, and could be able to distinguish between the primary tissue of relapsing and non-relapsing patients at the protein level.

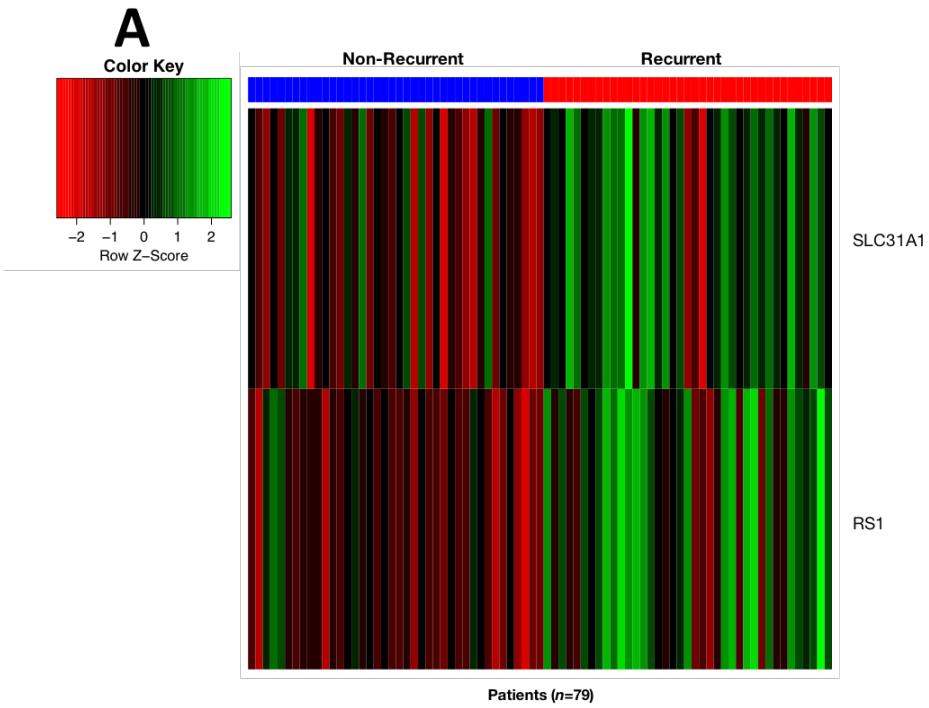
Interestingly, differential gene expression analysis revealed that only two genes were significantly differentially expressed (corrected p-value<0.05) between recurrent and non-recurrent patient tissues in this dataset. The differentially expressed genes were SLC31A1, a copper transporter, and RS1, an extracellular matrix protein (Table 3.2), both of which were expressed more highly in recurrent than non-recurrent patient samples (Figure 3.2A). A member of the melanoma antigen family, MAGEA10, was expressed at lower levels in recurrent than non-recurrent patient samples, but had a p-value marginally higher than the significance threshold (p=0.0559) and was excluded from the list of candidates.

The bioinformatics approach provided two additional candidate biomarkers to investigate, and while testing antibodies on patient tissue (Chapter 4), further bioinformatic characterisation was performed on the more experimentally successful candidate RS1. The cBioportal (http://www.cbioportal.org/), a publically available toolset for performing meta-analysis on multi-dimensional microarray datasets in cancer, was used to investigate the mutation status of the RS1 locus in ten large prostate cancer microarray datasets. Mutations in the coding sequence of RS1 were rarely observed in prostate tumours, no matter the subtype (Figure 3.2B). Amplification of the RS1 locus occurred at a low frequency in localised prostate tumours (1.5 and 2.1% of cases in TCGA datasets) with four of six primary tumour datasets having no detected alterations at the RS1 locus. However, RS1 amplification was a considerably more

frequent phenomenon in metastatic and castration resistant prostate tumours, with between 7.3% and 31.8% of patients in these four metastatic/CRPC datasets having altered copy number at the locus. These data suggest that RS1 expression could be altered during tumour progression at the level of copy number, and the findings add more evidence for the briefly-studied protein being a useful candidate biomarker.

Table 3.2: Top 10 differentially expressed genes between recurrent and non-recurrent prostate tumours in the Sun and Goodison dataset (GDS4109).

Probe ID	Gene Name	Gene Symbol	Chromosomal Location	Log_2 Fold-Change	P-Value
216937_s_at	retinoschisin 1	RS1	Xp22.13	0.551621798	0.004651396*
203971_at	solute carrier family 31 (copper transporter), member 1	SLC31A1	9q32	0.495047908	0.018742401*
210295_at	melanoma antigen family A, 10	MAGEA10	Xq28	-0.733052962	0.055949495
207683_at	forkhead box N1	FOXN1	17q11-q12	-0.755986662	0.078262367
203095_at	mitochondrial translational initiation factor 2	MTIF2	2p16.1	-0.898023549	0.078262367
215285_s_at	putative homeodomain transcription factor 1	PHTF1	1p13	0.316130005	0.101162106
207111_at	egf-like module containing, mucin-like, hormone receptor-like 1	EMR1	19p13.3	0.440721739	0.101162106
214146_s_at	pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)	PPBP	4q12-q13	0.752298133	0.101162106
221969_at	paired box 5	PAX5	9p13	-0.508480391	0.101162106
208099_x_at	tubulin tyrosine ligase-like family, member 5	TTLL5	14q24.3	-0.644293208	0.101162106


A
Color Key
−2 −1 0 1 2
Row Z−Score
Non-Recurrent
Recurrent
SLC31A1
RS1
Patients (n=79)

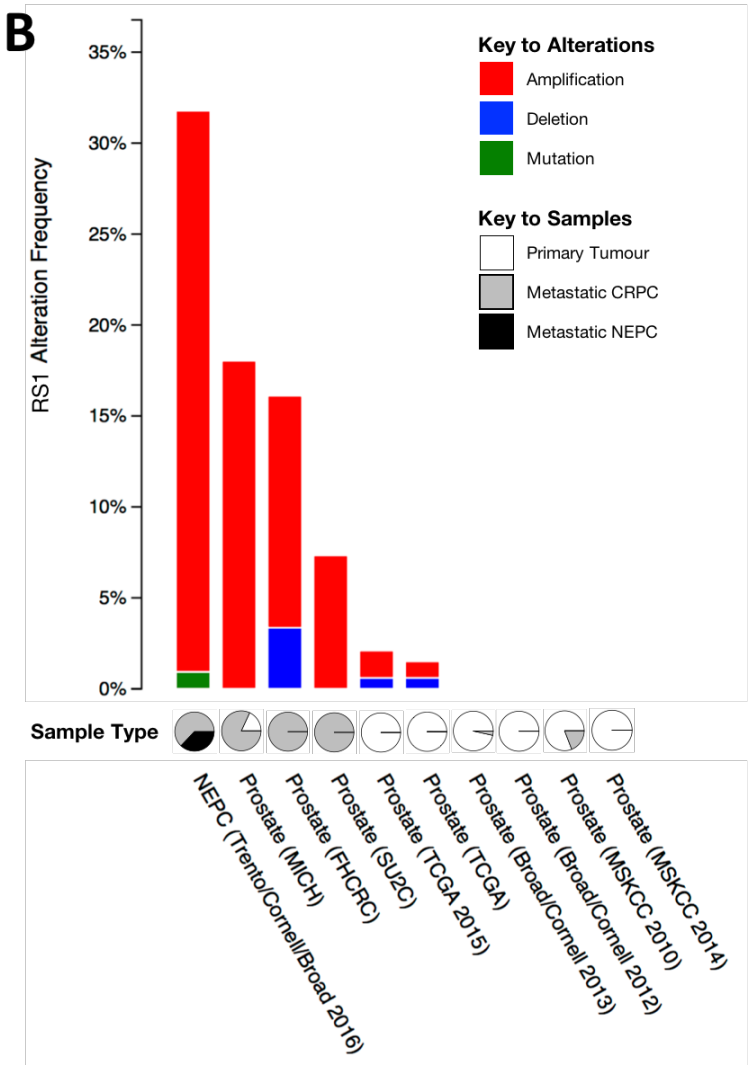

B
RS1 Alteration Frequency
35%
30%
25%
20%
15%
10%
5%
0%
Key to Alterations
Amplification
Deletion
Mutation
Key to Samples
Primary Tumour
Metastatic CRPC
Metastatic NEPC
Sample Type
NEPC (Trento/Cornell/Broad 2016)
Prostate (MICH)
Prostate (FHCRC)
Prostate (SU2C)
Prostate (TCGA 2015)
Prostate (TCGA)
Prostate (Broad/Cornell 2013)
Prostate (Broad/Cornell 2012)
Prostate (MSKCC 2010)
Prostate (MSKCC 2014)

Figure 3.2: Differential gene expression analysis based on recurrence status from Sun and Goodison (GDS4109) and subsequent meta-analysis of the identified biomarker RS1.

(A) Differential gene expression analysis performed on the Sun and Goodison (2009) dataset, showing the two significant ($P<0.05$) differentially expressed genes between the non-recurrent (n=40, blue labels) and recurrent (n=39, red labels) patient groups. Z-scores are plotted with red and green indicating lower and higher than average expression in the recurrent group, respectively. (B) Stacked bar chart showing a cBioportal meta-analysis of 10 prostate cancer microarray datasets. The frequency of amplifications, deletions and mutations of the RS1 locus in each dataset is shown as a percentage of the total cases, with the type and proportion of samples shown in a pie chart underneath: primary tumours; metastatic castration resistant prostate carcinomas (CRPC); or metastatic neuroendocrine prostate tumours (NEPC). Each bar is labelled with a study identifier as it is known in cBioportal.

3.2.3 Estimation of biomarker functions through bioinformatics

3.2.3.1 Extremes of expression analysis

Having identified candidate biomarkers to test using two different methods, I first wanted to learn more about the functions of these biomarkers in relevant cell types, particularly for the less well-characterised biomarkers such as ZSCAN4, where little or no information may be known about its relevance to tumour biology and the functions it performs. To do this, I developed a new method of analysis using existing large microarray datasets. By segregating the samples in each dataset to correspond to the highest and lowest 10% of intensity values for a given probe on the array, discarding the samples in between and comparing the two groups, I hypothesised that it would be possible to learn more about the gene expression environment that arises in the context of elevated biomarker expression levels. Thus, the approach is a differential gene expression based analysis using a gene's expression levels as a grouping system, with exclusion of the intermediate values providing greater contrast to the comparison. Full code for this method can be found in Appendix 1.

Differentially expressed genes can be fed into a gene ontology (GO) annotation package and visualised, in this case using REVIGO (http://revigo.irb.hr/) meaning that gene functions that are overrepresented as a consequence of high biomarker expression can

be found and visualised (Supek *et al.*, 2011). These ontologies do not represent the specific function of the biomarker, but rather the collective functions of the genes whose expression changes as a result of changes in biomarker expression. REVIGO takes gene ontology categories and groups them using semantic similarity methods, meaning that similar gene ontology terms are clustered together in space. For ease of visualisation and interpretation, only the top 20 gene ontology terms are used in this analysis. It is expected that these gene ontologies will cluster together in groups that are linked to the function of the biomarker itself within the context of the tissues or cells being studied. The aim was to study biomarkers of interest within prostate cancer datasets.

3.2.3.2 Method validation

To validate this method of analysis, a gene with known functions was investigated, and the 10^{th} and 90^{th} percentiles of expression were compared to assess functional categories of the differentially expressed genes. Two datasets were chosen for this analysis. The first was a single platform of the Grasso dataset (GEO accession: GSE35988, platform GPL6480) which consists of 88 patients: 49 primary prostate tumours, 27 metastatic castration resistant prostate tumours and 12 benign prostate tissue samples. The second was the larger CamCaP dataset (GEO accession: GSE70768) which consists of 199 primary prostate samples: 125 prostate tumours and 74 matched benign prostate tissue samples. Aurora A Kinase (AURKA) was chosen as a control, given its well-established function as a mitotic kinase involved in the regulation of multiple mitosis checkpoints, including chromosomal alignment and spindle assembly. Therefore, differentially expressed genes between samples with high and low AURKA expression would be expected to include those with roles in mitosis, and gene ontology terms should be consistent with this.

In agreement with this hypothesis, upregulated genes in samples with a high AURKA expression were significantly enriched for gene ontology biological process terms to do with the cell cycle, cell division, chromosome segregation and protein complex disassembly among others (Figure 3.2A and B), when compared to samples with low AURKA expression. This finding is consistent among two large microarray datasets of prostate cancer patient tissue (Table 3.2) and is also consistent with the known functions of AURKA in the control of the mitotic cell cycle and spindle assembly checkpoint (Nikonova *et al.*, 2013), demonstrating that the extremes of expression method works for this gene in several prostate cancer datasets.

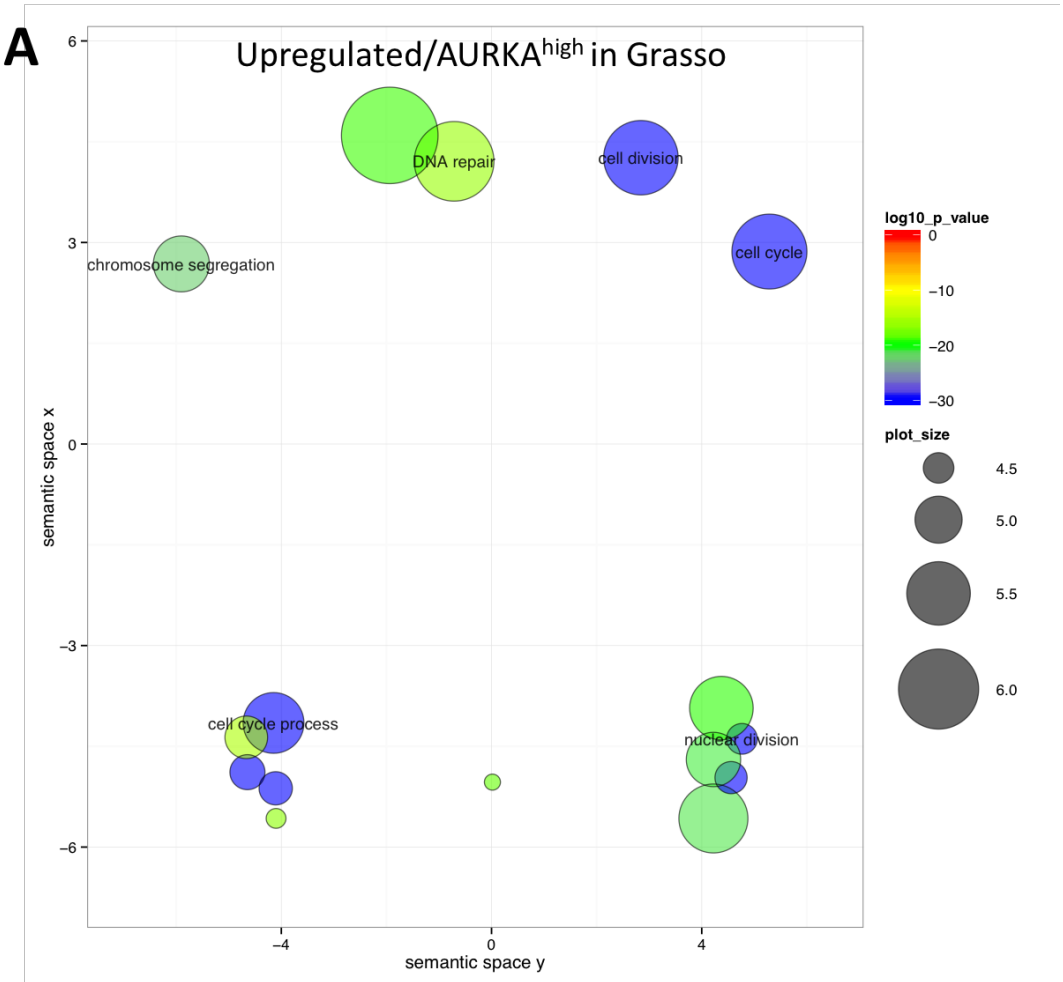

A
Upregulated/AURKA^high in Grasso
DNA repair
cell division
cell cycle
chromosome segregation
cell cycle process
nuclear division
semantic space x
semantic space y
log10_p_value
plot_size
B
Upregulated/AURKA^high in CamCaP
translation
nuclear−transcribed mRNA catabolic process, nonsense−mediated decay
protein complex disassembly
cellular metabolic process
cell division
cell cycle
mitotic cell cycle
semantic space x
semantic space y
log10_p_value
plot_size

Figure 3.3: Gene ontology (GO) classification of upregulated genes in prostate cancer patients with high AURKA expression in two independent cohorts.

(A) Visual representation of GO terms enriched in patients with high AURKA expression in the Grasso cohort (GSE35988). GO terms are grouped together in semantic space using REVIGO, where terms group together into those involved in the cell cycle, DNA repair and nuclear division as expected. (B) Visual representation of GO terms enriched in patients with high AURKA expression in the CamCaP cohort (GSE70768), where terms involving the cell cycle, cell division and protein complex disassembly are represented, along with translation.

Table 3.3: Top 5 significantly enriched gene ontology terms when AURKA expression is high or low, in two large prostate cancer patient cohorts.

Cohort	*Rank*	*Upregulated in AURKA^{high}*				
		GO Term	**Annotated**	**Signif.**	**Expect.**	**p-val**
Grasso	1	cell cycle	1564	93	21.04	< 1e-30
	2	mitotic cell cycle	923	72	12.42	< 1e-30
	3	cell cycle process	1161	79	15.62	< 1e-30
	4	mitotic cell cycle process	793	67	10.67	< 1e-30
	5	nuclear division	495	55	6.66	< 1e-30
Cam CaP	1	mitotic cell cycle	927	48	17.24	8.20E-11
	2	translational termination	176	20	3.27	1.00E-10
	3	translational elongation	205	21	3.81	2.40E-10
	4	cellular protein complex disassembly	273	24	5.08	2.90E-10
	5	protein complex disassembly	297	25	5.52	3.10E-10

3.2.3.3 Investigation of SDC1 in human transcriptomic datasets

Having validated the extremes of expression methodology, I used the method to carry out investigations of our candidate markers. Only markers for which preliminary data was promising in initial human testing (presented in the next two chapters) were investigated – RS1, SDC1 and ZSCAN4. Unfortunately, intensity values of RS1 probes on the Grasso and CamCaP datasets were too low and many samples had undetectable expression of RS1 in the probesets, which excluded RS1 from this analysis method. I began by identifying the gene ontology terms associated with the expression of SDC1. As before, both datasets were separated into their 10^{th} and 90^{th} percentiles, with extremes compared and intermediate samples removed from the analysis.

When extremes of expression analysis was applied to study the effects of varying SDC1 expression in the context of prostate cancer datasets (Grasso and CamCaP), the prominent GO terms represented were different between both datasets. In Grasso, high expression of SDC1 was associated with an enrichment of gene expression involved in chromosome and organelle organisation and segregation (Figure 3.4A). 125 genes were also annotated to the acute immune response (Table 3.3). In the CamCaP dataset, the high expression of SDC1 was not associated with chromosome organisation or immune response GO terms, but instead with genes involved in epithelial morphogenesis, epithelial development, cell motility and cell adhesion (Figure 3.4B). Interestingly, while SDC1 has potential roles in cell adhesion, development, motility and the immune response, a relationship between SDC1 expression and chromosome organisation has not been previously established. The generation of different GO enrichments for both datasets indicates that there is considerable variability between individual microarray datasets.

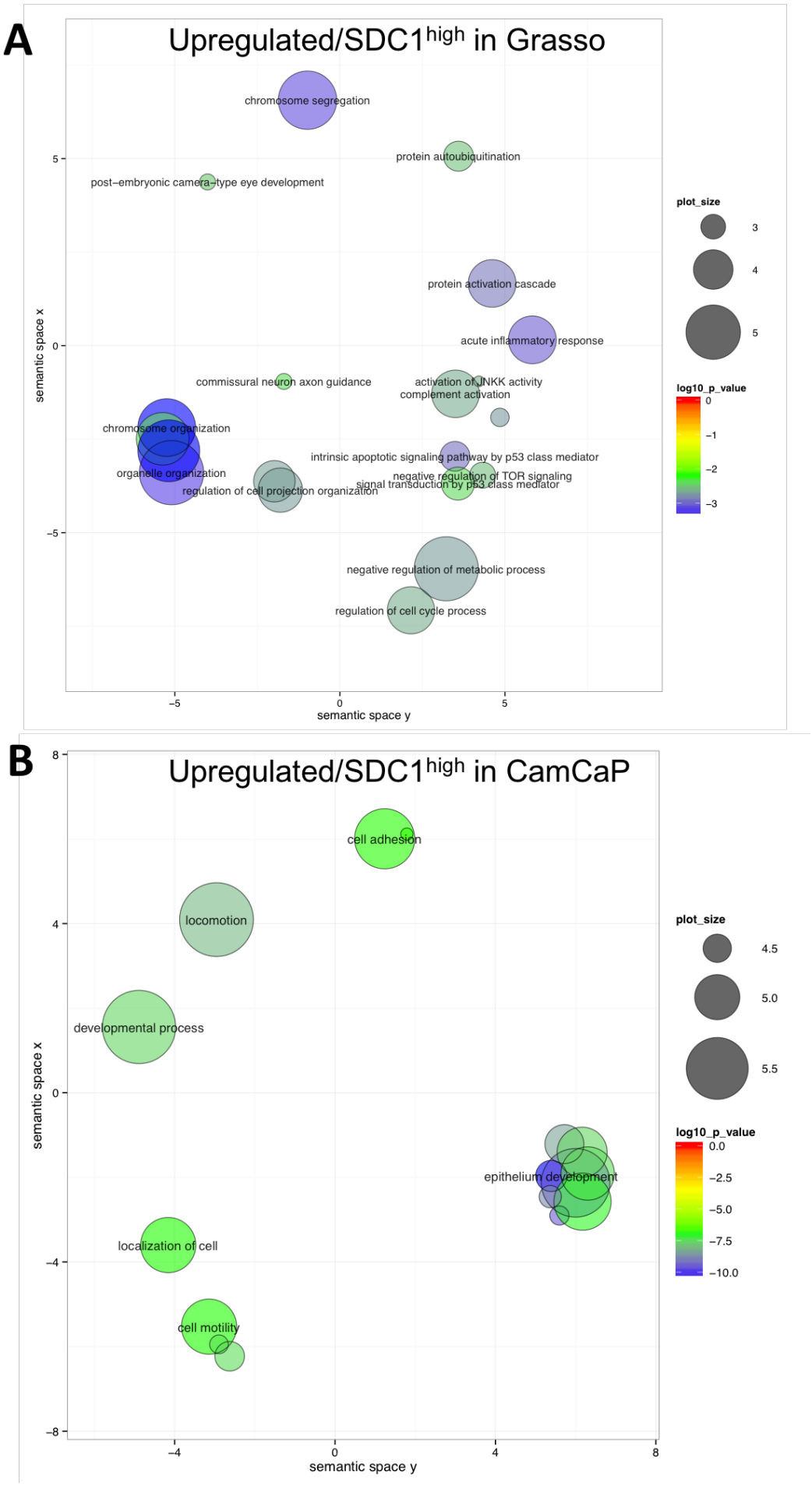

A
Upregulated/SDC1high in Grasso
chromosome segregation
protein autoubiquitination
post−embryonic camera−type eye development
protein activation cascade
acute inflammatory response
commissural neuron axon guidance
activation of JNKK activity
complement activation
chromosome organization
organelle organization
regulation of cell projection organization
intrinsic apoptotic signaling pathway by p53 class mediator
negative regulation of TOR signaling
signal transduction by p53 class mediator
negative regulation of metabolic process
regulation of cell cycle process
semantic space x
semantic space y
plot_size
3
4
5
log10_p_value
0
−1
−2
−3
B
Upregulated/SDC1high in CamCaP
cell adhesion
locomotion
developmental process
epithelium development
localization of cell
cell motility
semantic space x
semantic space y
plot_size
4.5
5.0
5.5
log10_p_value
0.0
−2.5
−5.0
−7.5
−10.0

Figure 3.4: Gene ontology classification of upregulated genes in prostate cancer patients with high or low SDC1 expression in the CamCaP Cohort.

(A) Visual representation of GO terms enriched in patients with high SDC1 expression in the Grasso cohort (GSE35988). GO terms are grouped together in semantic space using REVIGO, showing a diverse range of terms concerning chromosome organisation, the acute inflammatory response, complement activation, cell cycle, metabolism and cell projection. (B) Visual representation of GO terms enriched in patients with high SDC1 expression in the CamCaP cohort (GSE70768), where terms involving the epithelial development, cell adhesion and motility are represented.

Table 3.4: Top 5 significantly enriched gene ontology terms when SDC1 expression is high, in two large prostate cancer patient cohorts: Grasso (GSE35988) and CamCaP (GSE70768).

Cohort	*Rank*	*Upregulated in SDC1^{high}*				
		GO Term	**Annotated**	**Signif.**	**Expect.**	**p-value**
Grasso	1	chromosome organization	931	24	11.45	0.00046
	2	single-organism organelle organization	2242	45	27.57	0.00046
	3	organelle organization	3178	58	39.08	0.00066
	4	chromosome segregation	239	10	2.94	0.00076
	5	acute inflammatory response	125	7	1.54	0.00089
Cam CaP	1	epithelium development	1054	57	22.36	4.20E-11
	2	anatomical structure morphogenesis	2465	97	52.28	2.30E-10
	3	morphogenesis of an epithelium	466	34	9.88	3.30E-10
	4	tissue morphogenesis	600	38	12.73	1.60E-09
	5	anatomical structure development	4754	151	100.83	2.20E-09

3.2.3.4 Investigation of ZSCAN4 expression in human transcriptomic datasets

Given that very little was known about ZSCAN4 expression in human tissues prior to this study, I first investigated where ZSCAN4 might be expressed using high throughput sequencing data from the ENCODE project (Dunham *et al.*, 2012), using the Integrative Genomics Viewer to visualise it (Robinson *et al.*, 2011; Thorvaldsdottir *et al.*, 2013). ENCODE Chip-Seq data was visualised along with Illumina Bodymap 2.0 RNA-Seq data for a variety of normal human tissues and cell lines and it was found that placenta tissue expressed the ZSCAN4 transcript (Figure 3.5). Prostate tissue expressed the transcript to a much lower level. However, in both cases it was found that the cDNA sequencing reads extended much further to the 5' of the ZSCAN4 ORF than would be expected, with reads in both placenta and prostate tissues being detected as far as approximately 17kb upstream of the ZSCAN4 ORF (Figure 3.5, arrows), with a small Pseudogene located just downstream of this location. This raised the hypothesis that there might be multiple transcription start sites for the ZSCAN4 gene, with one approximately 4.5kb upstream of the ZSCAN4 ORF and the other approximately 17kb upstream.

In attempt to confirm this finding, information was gathered from normal human epidermal keratinocyte (NHEK) cells, which had ENCODE Chip-Seq data, to look for histone marks indicative of transcription start sites in the region including the activating histone marks H3k9ac, H3k4me3 and H3k27ac and repressing histone mark H3k27me3. Additionally, Cap Analysis Gene Expression (CAGE) data, where the 5' ends of transcripts are captured and sequenced, was taken from normal prostate tissue and overlayed onto RNA-Seq and Chip-Seq data. Clear peaks were observed corresponding to activating histone marks in NHEK cells at both hypothetical transcription start sites, with prostate CAGE data also showing peaks of sequenced 5' transcripts at these sites (Figure 3.5, arrows). Repressive histone marks were absent. Taken together, the data suggest that ZSCAN4 is expressed in the human placenta, which could serve as a possible positive control tissue for later use. It also suggests that there may be longer ZSCAN4 transcripts than previously documented, with possible alternative transcriptional start sites, although this would require further experimental validation.

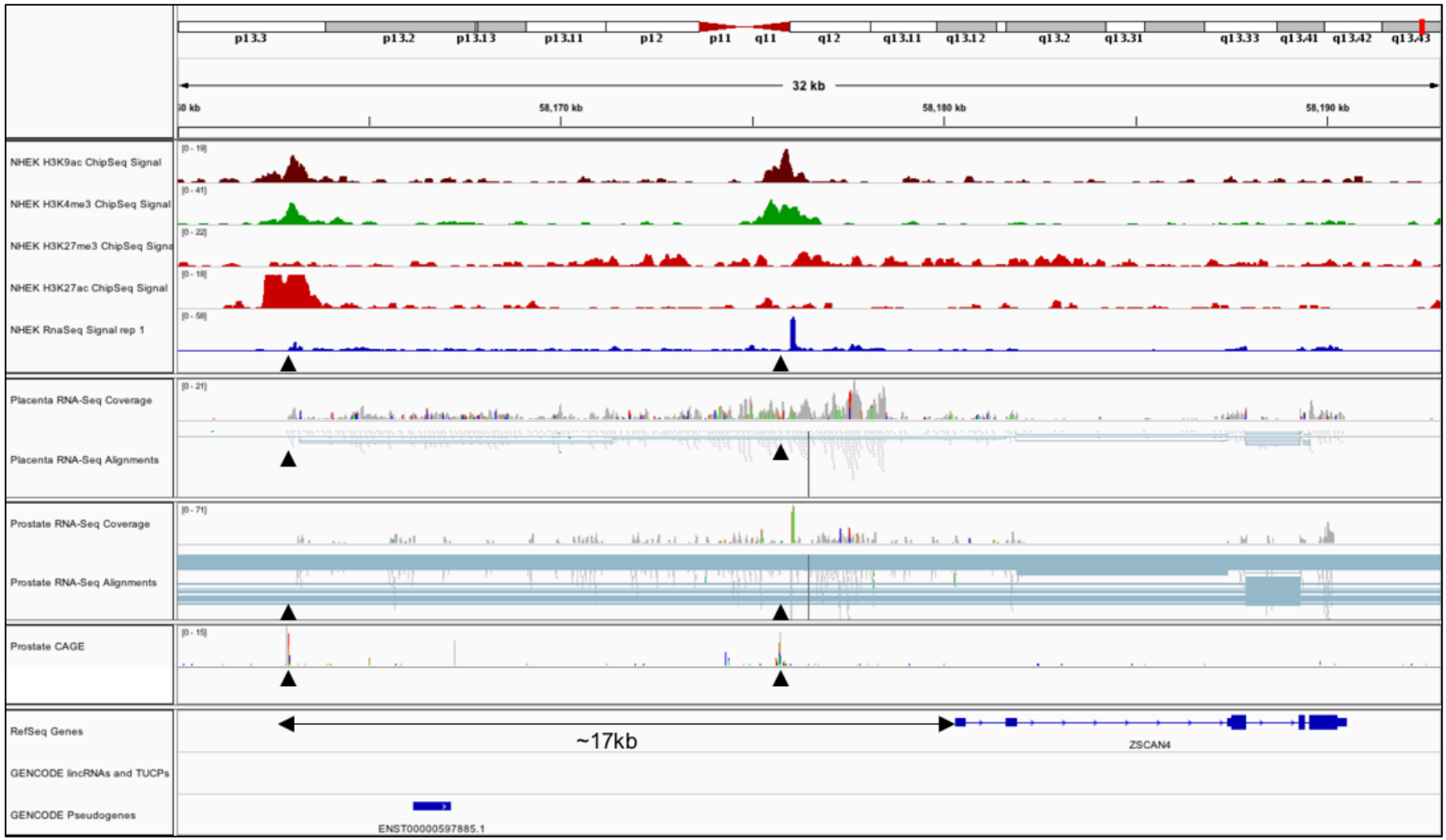

p13.3
p13.2
p13.13
p13.11
p12
p11
q11
q12
q13.11
q13.12
q13.2
q13.31
q13.33
q13.41
q13.42
q13.43
32 kb
58,170 kb
58,180 kb
58,190 kb
NHEK H3K9ac ChipSeq Signal
NHEK H3K4me3 ChipSeq Signal
NHEK H3K27me3 ChipSeq Signa
NHEK H3K27ac ChipSeq Signal
NHEK RnaSeq Signal rep 1
Placenta RNA-Seq Coverage
Placenta RNA-Seq Alignments
Prostate RNA-Seq Coverage
Prostate RNA-Seq Alignments
Prostate CAGE
RefSeq Genes
GENCODE lincRNAs and TUCPs
GENCODE Pseudogenes
~17kb
ZSCAN4
ENST00000597885.1

Figure 3.5: Chip-Seq and RNA-Seq data from human cells and tissues in the ENCODE and Illumina Bodymap 2.0 projects, showing the ZSCAN4 locus and the region immediately upstream.

Refseq data shows the genomic region containing the ZSCAN4 transcript (blue, bottom panel), whilst two regions approximately 4.5kb and 17kb away from the annotated ZSCAN4 transcript have regions of histone methylation consistent with active transcription start sites (arrowheads), such as H3k9ac, H3k4me3 and H3k27ac as determined by Chip-Seq. Repressive chromatin marks such as H3k27me3 are absent. Similarly, CAGE data from prostate tissue captures the 5' end of transcripts at both TSS-like areas (arrowheads). While prostate tissue has low expression of ZSCAN4, placenta tissue shows evidence of more complete ZSCAN4 transcripts as determined by RNA-Seq. Transcripts are present from the -17kb position to the end of the last ZSCAN4 exon. GENCODE data reveals a small pseudogene mapped downstream of the -17kb putative transcription start site. Data were visualised in Integrative Genomics Viewer.

In order to interrogate the potential functions that ZSCAN4 might be having in tissues where it is expressed, I first identified the effects that variations in its expression would have in a tissue that we predicted would express ZSCAN4, based on transcriptomics data. ZSCAN4 probe intensity values were often very low or absent, making it difficult to assess the function of ZSCAN4 expression in the Grasso and CamCaP datasets. Instead, we used the human placenta microarray dataset from Bruchova et al. (2010) which contained 76 samples (GEO accession GDS3793) and performed extremes of expression analysis on it based on extremes of ZSCAN4 expression levels. A placental dataset was chosen because of the prior identification of ZSCAN4 transcripts expressed in placental RNA-seq datasets (Figure 3.5). In samples with high ZSCAN4 expression, GO terms involving metabolism and intracellular transport were overrepresented (Table 3.5 and Figure 3.6). Although comparisons with the literature are difficult given the lack of documentation of ZSCAN4 expression in placenta, together with only brief characterisation outside of ES cells, this analysis reveals a hypothetical role for ZSCAN4 in metabolism, and it would be interesting to see if future functional analysis of ZSCAN4 reveals such roles.

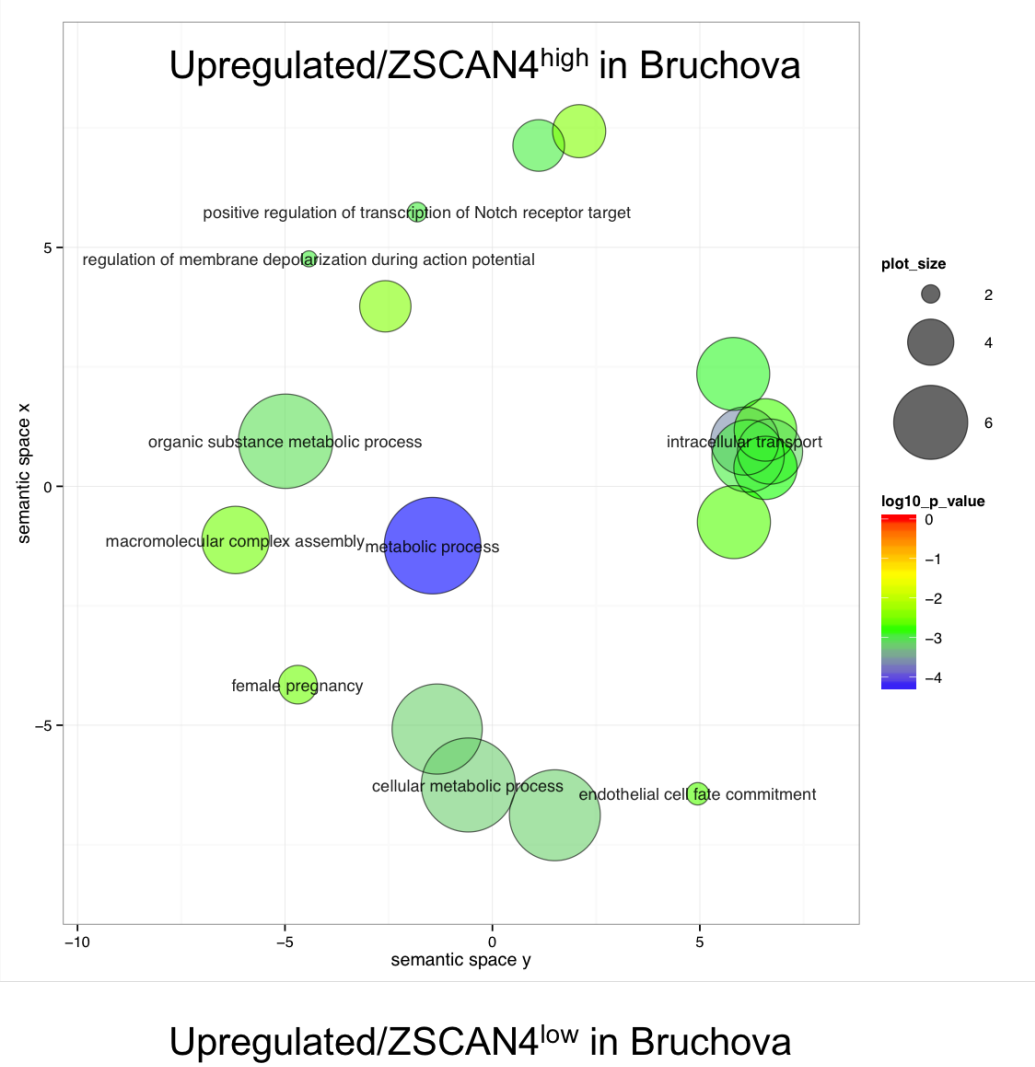


Figure 3.6: Gene ontology classification of upregulated genes in normal placenta tissue samples with high ZSCAN4 expression in the Bruchova Cohort.

Visual representation of GO terms enriched in patients with high ZSCAN4 expression in the Bruchova cohort (GDS3793). GO terms are grouped together in semantic space using REVIGO, with the majority of terms relating to metabolism and intracellular transport, with pregnancy, Notch signalling, endothelial cell fate, and other diverse terms also represented.

Table 3.5: Top 10 significantly enriched gene ontology terms when ZSCAN4 expression is high in Bruchova's large placenta dataset (GDS3793).

Cohort	*Rank*	*Upregulated in ZSCAN4^{high}*				
		GO Term	Annotated	Signif.	Expect.	p-value
Bruchova	1	metabolic process	10633	567	521.46	5.30E-05
	2	intracellular transport	1642	111	80.53	0.00025
	3	cellular metabolic process	9066	486	444.61	0.00066
	4	macromolecule metabolic process	7841	427	384.53	0.00068
	5	cellular macromolecule metabolic process	7186	395	352.41	0.00069
	6	single-organism intracellular transport	1338	91	65.62	0.00078
	7	organic substance metabolic process	9277	495	454.96	0.00088
	8	establishment of localization in cell	2350	146	115.25	0.00109
	9	asparagine biosynthetic process	5	3	0.25	0.00109
	10	positive regulation of transcription of Notch receptor target	5	3	0.25	0.00109

3.2.4 Optimisation of anti-biomarker antibodies for use in IHC

When the potential biomarkers had been selected by both approaches and characterisation had been done at the bioinformatic level, the next step was to establish and test methods for the IHC procedure to be used on human samples. Since the initial aim was to construct a tissue microarray of precious material such as prostate biopsies and prostatectomy samples, it was important to test a method for constructing tissue microarrays from FFPE tissue samples. In addition, it was necessary to ensure that IHC procedures would work on the constructed tissue microarray, and some antibodies raised against candidate markers could be tested in this setting. Therefore, I tested a method for constructing and staining tissue microarrays using a range of normal FFPE tissue from mice.

3.2.4.1 Development of a mouse tissue microarray for testing

Mouse tissue is a readily available source of control material which does not require extended ethical consideration to obtain, and tissue preparation can be controlled and standardised by the end user. This made it suitable for construction of a trial tissue microarray in which construction and staining protocols could be tested. Additionally, a wide range of different tissue types could be harvested from the mice so that a multi-tissue survey could be conducted on this tissue microarray, and archived mouse FFPE tissue blocks that were already prepared by Dr Asha Recino in 2007 were used for most of the source cores, with the exception of prostate tissue which was prepared freshly for the purpose of producing the TMA.

H&E staining of whole sections from blocks of mouse tissue was carried out, and tissue microarrays were constructed using the whole sections to guide core punch placement. This resulted in a tissue microarray containing 48 cores of mouse tissue from 13 different organs (Figure 3.7), making it a useful resource for surveying protein expression patterns in the mouse. The tissue microarray was arranged into a defined, asymmetrical pattern of tissue cores divided between three sectors (Figure 3.7) to ensure consistent orientation and simplify navigating the array under a microscope. Sectors were divided by a row of blank paraffin cores and the array was lined by a single layer of blank paraffin cores to aid in sectioning. 8 cores of prostate tissue were included in order to describe the heterogeneous tissues present in the four murine prostate lobes.

Figure 3.7: Schematic of tissue microarray cores in the mouse tissue microarray.

Tissues are colour keyed, numbered by specimen number and gender denoted where known. The design is asymmetric, with an orientation core of dyed agarose placed in the corner to assist in microarray orientation during sectioning.

H&E staining performed on the tissue microarray confirmed that cellular morphology of organs was intact after microarray construction: liver (Figure 3.8A); kidney (Figure 3.8B); lung (Figure 3.8C); intestine (Figure 3.8D); brain (Figure 3.8E); skin (Figure 3.8F); testis (Figure 3.8G); pancreas (Figure 3.8H); and skeletal muscle (Figure 3.8I) were all consistently retained in the tissue microarray. In addition, some tissues too brittle to easily section whole, such as liver and skeletal muscle (Figures 3.8A & I), were successfully retained in microarray sections. From a single section of tissue microarray, morphological features of three prostate lobes were apparent (Figure 3.9). Following H&E staining, prostate cores retained gross glandular architecture and the surrounding fibromuscular stroma (Figure 3.9A). Ventral prostate glands displayed characteristic focal tufting and cuboidal epithelium (Figure 3.9B, arrow), while anterior prostate showed glandular infolding (Figure 3.9C, arrow) and prominent eosinophilic secretions. The epithelium of the dorsolateral prostate (Figure 3.9D) was cuboidal and surrounded by a loose connective tissue.

Tissue types where available tissue was scant, such as ovary and spleen, showed frequent core loss when IHC was conducted. Over 21 IHC staining reactions performed on sections of the tissue microarray, the mean percentage of cores lost was 29.7%. As the majority of tissue types were intact in this array, this high rate of core loss was considered acceptable for the initial IHC studies.

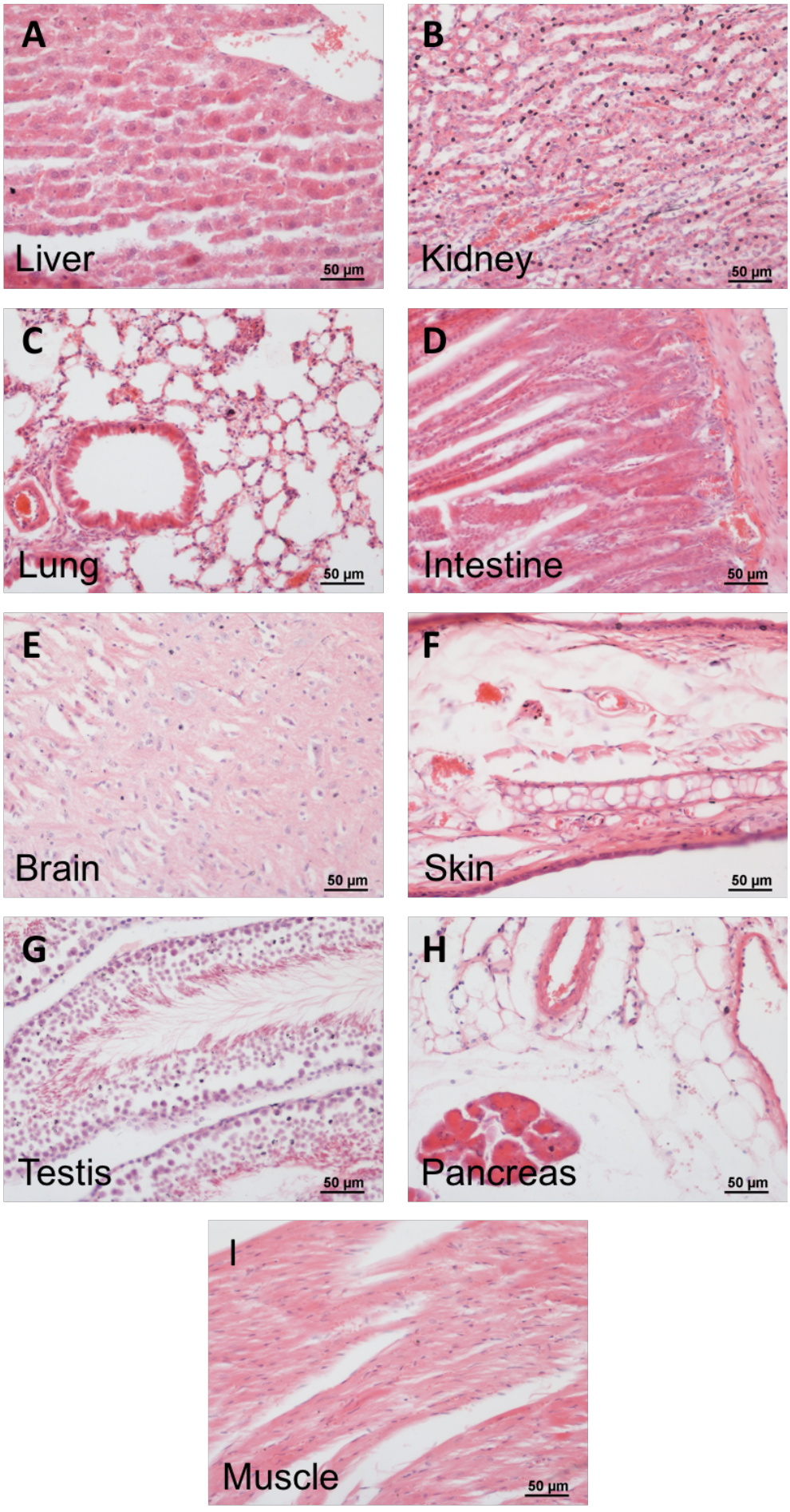

A
Liver
50 µm
B
Kidney
50 µm
C
Lung
50 µm
D
Intestine
50 µm
E
Brain
50 µm
F
Skin
50 µm
G
Testis
50 µm
H
Pancreas
50 µm
I
Muscle
50 µm

Figure 3.8: H&E staining of representative mouse tissue microarray cores at 20x magnification.

Images of H&E stained cores of a single mouse tissue microarray section: (A) Liver; (B) Kidney; (C) Lung; (D) Intestine; (E) Brain; (F) Skin; (G) Testis; (H) Pancreas; (I) Skeletal muscle.

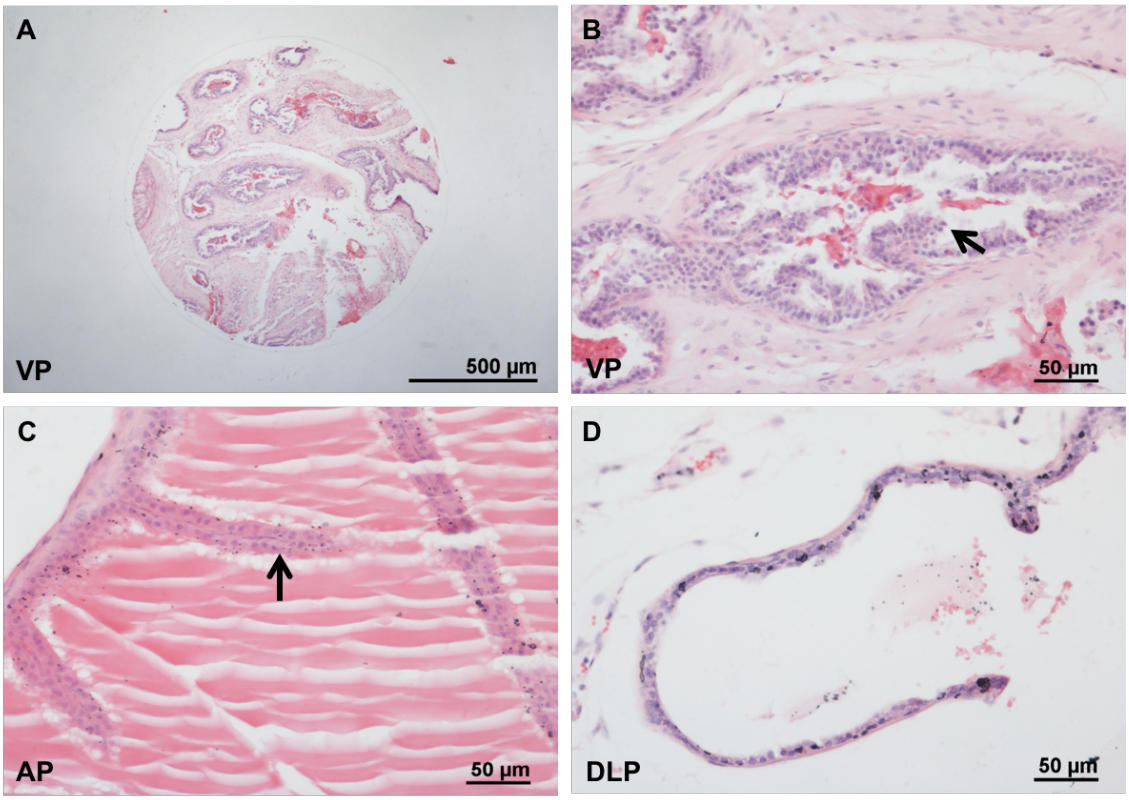


Figure 3.9: H&E staining of mouse prostate tissue microarray cores viewed at high and low magnification.

(A) Low power view of a whole core of ventral prostate, demonstrating that the glandular and stromal architecture remains intact after sectioning. (B) High power view of a murine ventral prostate gland as in (A) demonstrating focal tufting (arrow). (C) Anterior prostate with glandular infolding (arrow) and prominent eosinophilic secretions. (D) Dorsolateral prostate with small nuclei and compact, cuboidal epithelium. Scale bars – (A) 500µm, (B-D) 50µm. VP – ventral prostate; AP – anterior prostate; DLP – dorsolateral prostate.

3.2.4.2 Establishing the mouse tissue microarray as an IHC testing platform

In order to demonstrate the suitability of the tissue microarray for downstream IHC analysis when applied to prostate tumour specimens, initial IHC staining was performed on the mouse tissue microarray and some candidate marker antibodies were tested in the process. Initial optimisation of IHC protocols was performed on two of six candidates identified in the literature: ZSCAN4 and ALDH7A1. Where known, expression patterns across mouse tissues were cross-checked with the results obtained in this study. The antibodies for the other four markers, BMI1, NES, SDC1 and MUC1-C, were raised against human antigens and not shown to identify the murine homologues in the literature, or found not to react with the mouse tissue microarray. For this reason, these four markers were not optimised using the mouse tissue microarray.

In addition to testing antibodies with the staining procedure, the level of non-specific, background staining was assessed by carrying out parallel negative controls in which the same secondary detection system was used, but the primary antibody was not applied. As with the test antibody slides, negative control slides were scored for their staining patterns where they were present. It was found that murine lung, kidney, spleen and heart tissue demonstrated diffuse cytoplasmic reactions even in the absence of primary antibodies, regardless of which species the secondary detection system was designed to detect (Appendix 2). However, the remaining tissues were largely clear of non-specific background signals and were suitable for assessment. Given this, it is likely that the background staining is a tissue-specific phenomenon, and therefore the array as a whole is suitable for assessment of antibody staining patterns provided that these background-prone tissues are taken into account.

3.2.4.3 ALDH7A1 staining was predominantly nuclear in most murine tissues

Expression of ALDH7A1 was detected in the nuclei of the majority of murine tissues stained, with considerable heterogeneity in intensity between individual cells. For example, a subset of hepatocytes showed stronger staining than surrounding cells (Figure 3.11A, arrow), while the murine pancreas contained a mixture of strongly positive and negatively staining nuclei (Figure 3.11B) and renal tubular epithelium had stronger nuclear staining than surrounding cells (Figure 3.11C). In contrast, prostate glands showed cytoplasmic staining of glandular epithelium with no nuclear staining (Figure 3.11D), or with weak to intermediate nuclear staining as shown in the lateral prostate (Figure 3.11E). Negative controls, where no primary antibody was added, confirm that the secondary detection systems are free of detectible background in

prostate specimens (Figure 3.11F). When the localisation of the ALDH7A1 staining pattern was scored, localisation was found to be nuclear within the vast majority of tissues over the course of three independent experimental replicates, with the exception of prostate where the predominant localisation was cytoplasmic (Appendix 3).

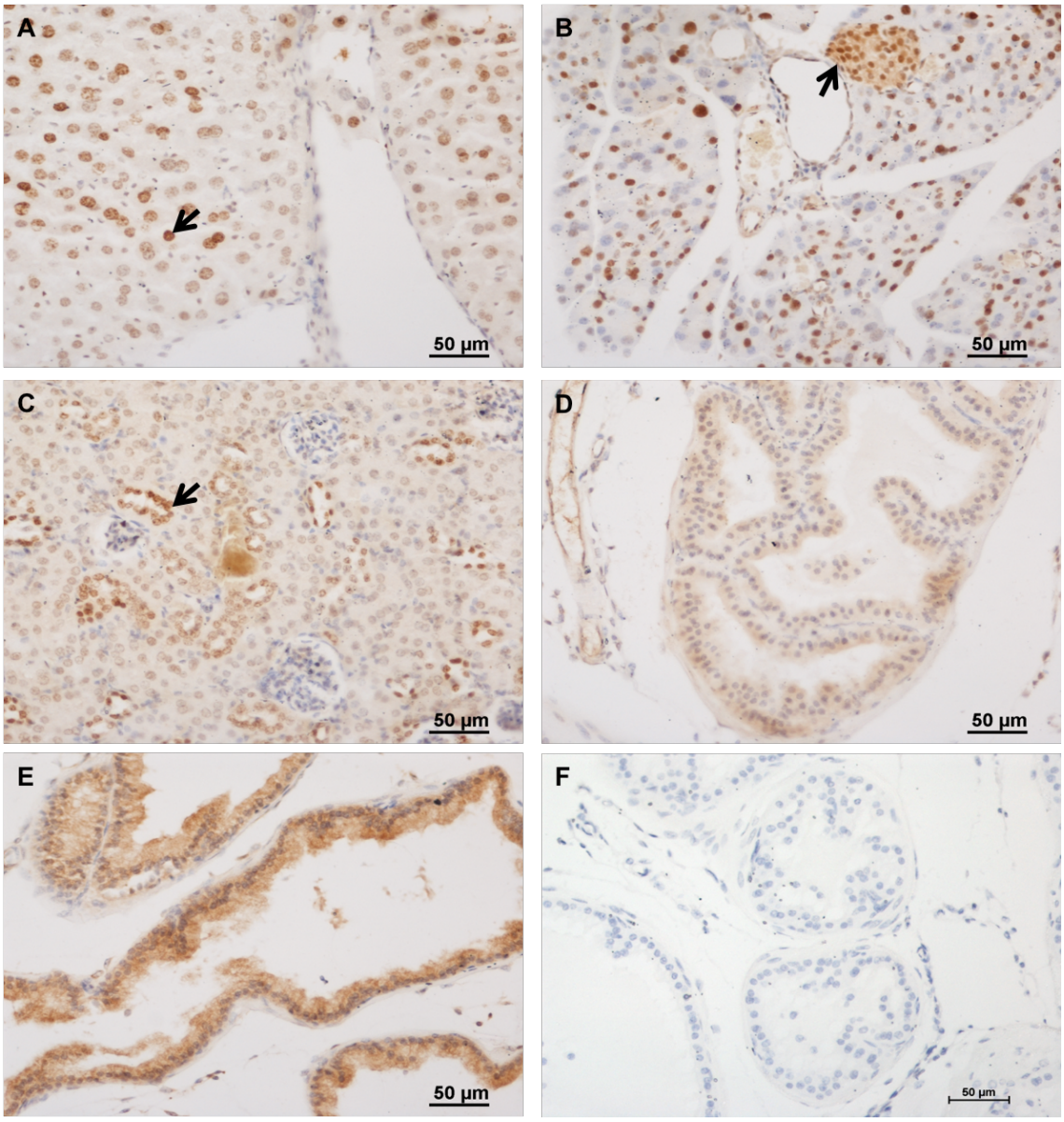


Figure 3.10: ALDH7A1 expression in adult mouse tissues assessed by IHC.

(A) Liver hepatocytes show weak to intermediate nuclear staining of ALDH7A1 (arrow). (B) Murine pancreas showing strong heterogeneous nuclear staining, with islets of Langerhans showing cytoplasmic and nuclear staining (arrow). (C) Kidney cells stain nuclear positive, with individual strongly staining tubules visible (arrow). (D) Prostates show uniform, weak to intermediate cytoplasmic staining of glandular epithelium, as seen in the anterior prostate and (E) lateral prostate. (F) Negative control material is anterior prostate tissue with no primary antibodies added. Images representative of 3 independent experiments. Scale bars – 50µm.

3.2.4.4 ZSCAN4 was diffuse and weak to negative in all murine tissues

Although expression of ZSCAN4 has been reported in several human tissue types at the transcript level, including placenta, lung, kidney, liver and pancreas (Ko *et al.*, 2013), its expression at the protein level in adult tissues has not been thoroughly studied and expression in the prostate has not been reported. In order to characterise its expression and optimise the antibody for downstream IHC staining protocols with precious biopsy tissue, we used immunohistochemistry on mouse tissue microarray sections. ZSCAN4 staining was detected at a low level in most tissues, with weak, diffuse cytoplasmic staining patterns being observed in all tissue types examined. Within the murine prostate, staining was weak to negative within the glandular epithelium (Figure 3.12A and B) with moderate staining of fibromuscular prostate tissue components (Figure 3.12B). Murine pancreas, a tissue previously demonstrated to contain a subset of ZSCAN4+ staining cells in all tissue compartments (Ko *et al.*, 2013), was diffusely and weakly positively stained with the antibody (Figure 3.12C), which does not match the published finding of Ko *et al.* (2013). Staining in the testis was limited to Leydig cells (Figure 3.12D, arrows), whereas in the liver staining was weak and diffuse across all hepatocytes (Figure 3.12E). Negative controls, where no primary antibody was added, demonstrate that the detection method has a low background (Figure 3.12F). However, the diffuse and weak/negative nature of the staining across all tissues (Appendix 4), as well as the lack of corroboration with previously published data, suggests that this antibody is ineffective for detecting ZSCAN4 protein of mouse origin.

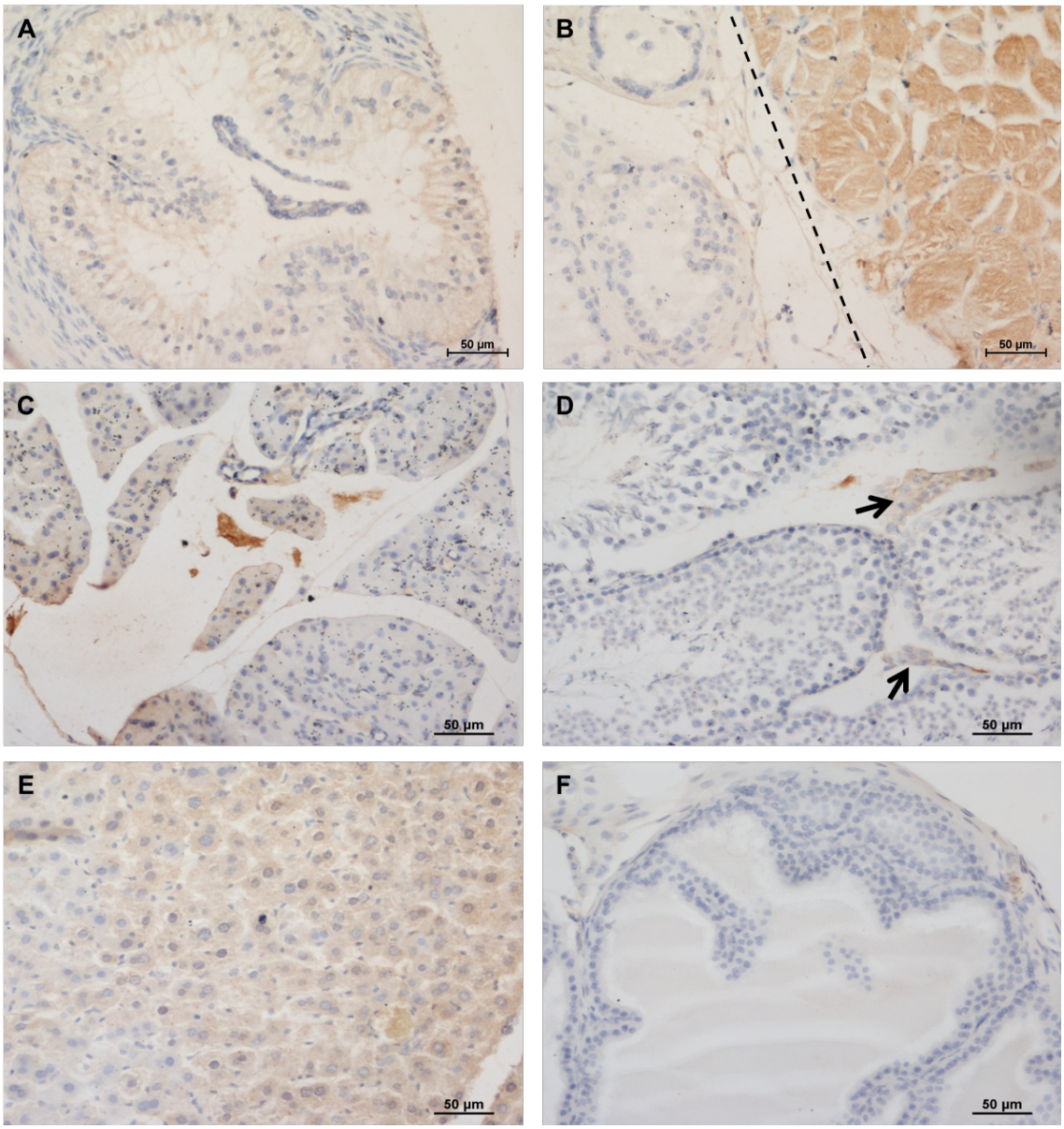

A
50 µm
B
50 µm
C
50 µm
D
50 µm
E
50 µm
F
50 µm

Figure 3.11: ZSCAN4 expression in adult mouse tissues.

(A) Prostate tissue shows weak diffuse glandular cytoplasmic staining, or (B) no staining in glandular epithelium, with a moderate intensity reaction in fibromuscular stroma (separated by dotted line). (C) Pancreatic tissue shows weak cytoplasmic staining. (D) ZSCAN4 expression in the testis is weak and cytoplasmic, being confined to Leydig cells (arrows). (E) Liver hepatocytes show weak, diffuse cytoplasmic staining. (F) Negative control is anterior prostate with no primary antibodies added. Figure 3.8A is representative of 2/3 experiments. Other images representative of 3/3 independent experiments. Scale bars - 50μm.

3.3 Discussion

In this chapter, potential biomarkers were selected using two different methods: six via a literature searching method and an additional two by a bioinformatics method using pre-existing cDNA microarray data from prostate cancer. A further bioinformatics approach was designed to interrogate the potential functions of biomarkers using cDNA microarray datasets. Following this, a tissue microarray was constructed using FFPE mouse tissue from a variety of tissue types. Using IHC it was confirmed that the platform was suitable for the technique, and some initial testing of anti-biomarker antibodies was performed.

3.3.1 Identification of biomarkers by literature review

The literature review approach generated a wide range of biomarkers which have been studied to differing degrees, ranging from candidates selected from transcriptomics screens to individually studied and well characterised oncogenes and cancer stem cell markers. Key cancer stem cell markers such as CD44 and CD133 were not chosen due to their already high degree of coverage and assessment in the literature (see Figure 3.1A), with the consideration that any biomarker utility would already have been demonstrated.

Since transcriptomics studies were common, it would have been possible to include many markers from a small number of such studies, but this approach was avoided in favour of concentrating on key markers that were clearly identified and focused on singly within the individual studies, especially with respect to prognosis or outcome. This is because the level of characterisation of individual markers in transcriptomics studies is minimal, and biomarkers derived from prognostic gene signatures may not have individual prognostic utility.

It is impossible to evaluate the efficacy of the search methods employed, but they have yielded over 116 targets for potential investigation by the laboratory. In hindsight, the use of a text-mining-based approach to complement the manual literature searching such as the SCAIView tool (Younesi *et al.*, 2012) would have provided additional benefits: reductions to the time spent searching, and increases to the number of candidates identified and the amount of evidence accrued for each biomarker. However, the approaches adopted in this project did yield 6 potential biomarkers for further analysis.

3.3.2 Identification of biomarkers by bioinformatics

Differential gene expression analysis identifies the genes that are significantly differentially expressed between two or more groups of samples. In the dataset studied, only two genes were discovered to be differentially expressed to significant levels – RS1 and SLC31A1 were overexpressed in patients that experienced recurrence within five years. The relatively small number of significant genes is interesting, as this suggests that when patients with and without a 5-year biochemical recurrence were very similar on the gene expression level. However, it could also be an indicator of heterogeneity in prostate tumours, a phenomenon which is already well known to impede attempts to create prognostic signatures using molecular profiles (Sboner *et al.*, 2010). Such variability is particularly prominent in higher risk prostate tumours (Wyatt *et al.*, 2014) and would reduce the resolving power of a differential gene expression analysis, making truly differentially expressed genes more difficult to detect.

A subsequent cBioportal meta-analysis identifies a subset of metastatic and castration resistant prostate carcinomas with amplification of the RS1 locus. Interestingly, this amplification was not seen in tumours of primary origin. RS1 is located at genomic location Xp22.13, and amplifications of the whole X chromosome are an observed event in prostate tumours (Visakorpi *et al.*, 1995), most notably in metastatic disease (Braun *et al.*, 2013). Therefore, it is unclear whether RS1 amplification is a consequence of this phenomenon and/or contributes a function in the context of prostate cancer progression. However, taken together with the increase in RS1 expression in recurrent primary tumours, this makes the candidate biomarker a more desirable target to study, and demonstrates that a combination of bioinformatic approaches can inform biomarker selection processes.

3.3.3 Extremes of expression analysis

3.3.3.1 Alternative methods

The extremes of expression method was designed to identify functional roles associated with genes of interest. It has been known for decades that there is a demonstrable tendency for genes with similar and coordinated expression patterns to show closely related functions (Eisen *et al.*, 1998), and the effect has recently been confirmed within gene ontology classifications in prostate cancer datasets (Jiang *et al.*, 2014). Therefore, clustering genes based on similarity of expression patterns over two or more experimental conditions is a common approach and it often makes use of correlation

coefficients to demonstrate co-expression (Reynier *et al.*, 2011). Functional annotation follows, with top-ranking genes in the correlation analysis being fed into a gene ontology annotation package. It has the advantage of working with smaller as well as with larger datasets, and could potentially combine data from multiple expression array platforms. However, while this method of analysis is a good test of similarity for expression patterns, correlation tests fail to take into account the magnitude of differences in gene expression over the range of expression of the gene of interest. Additionally, while there are more complex methods for functional analysis that take into account gene networks, copy number and mutation data within a prostate cancer context (Jiang *et al.*, 2014), they focus on a genome-wide analysis rather than focusing on a single gene of interest.

The extremes of expression method has less resolving power for gene expression pattern similarities, given that the data is dichotomised into two extreme groupings based on expression level of a gene of interest. However, this method tests for magnitudes of difference in gene expression levels between the two groupings, making it better able to resolve genes that are differentially expressed to a larger degree when the expression of the gene of interest changes. With further refinements it would be possible to adapt the method for a more robust multi-probe approach, by summing or averaging the intensity values from each probe corresponding to a gene of interest on an array, instead of relying on the readings from a single probe.

3.3.3.2 Method reliability

Using a gene with experimentally-validated and well characterised functions, AURKA, it was demonstrated that the extremes of expression method accurately identifies biological process annotations associated with the gene in the literature. Since the method has only been applied to transcriptomic datasets, it remains to be seen whether this method only works for genes that are regulated chiefly at the transcriptional level. For those that are regulated at the protein level, it would be interesting to see if such a method applied to an antibody array, which profiles the proteome using a matrix of different antibodies, would yield additional information. However, extremes of expression analysis of SDC1 also revealed known biological process annotations associated with the gene, whose biological activities are also regulated at the enzymatic level by cleavage to liberate a soluble ectodomain (Purushothaman *et al.*, 2010; Ramani and Sanderson, 2014). It suggests that the approach is suitable for the study regardless

of whether or not the protein directly regulates transcription as a transcription factor would.

3.3.3.3 SDC1

Many GO terms revealed by this analysis were consistent with known functions of SDC1, such as cell adhesion, cell migration, ECM organisation and regulation of immune response. In particular, endothelial SDC1 expression is important for regulating immune cell adhesion, with SDC1-knockout endothelium showing increased immune cell adhesion during shear stress (Voyvodic *et al.*, 2014), while SDC1-knockout macrophages exacerbated the inflammatory response and resulting damage associated with aortic aneurysm in another mouse model (Xiao *et al.*, 2012). These phenotypes suggest that SDC1 has a likely effect on inflammation that relies upon the interplay of cell adhesion and cell migration events.

GO terms such as epithelial morphogenesis and development were also strongly represented in $SDC1^{high}$ patients, suggesting that genes associated with a developmental programme are activated in the context of high SDC1 expression. This would be consistent with several reports in murine models, which suggest that some tissues require SDC1 for maintaining normal differentiation programme during development and in adult homeostasis. For example, SDC1 knockdown in murine neural progenitor cells resulted in premature differentiation and a reduction in beta-catenin levels in vivo (Wang *et al.*, 2012b), suggestive of an inability to respond to Wnt signalling. SDC1 homozygous knockout mice also suffer from an inability to respond to cold stress, and in this model SDC1 is required for adipocyte differentiation in vivo (Kasza *et al.*, 2014), and this effect is due to SDC1-mediated positive regulation of PPARγ signalling. Additionally, SDC1 expression is required for the maintenance of a putative cancer stem cell population in the prostate cancer cell line PC-3 (Shimada *et al.*, 2013). Given this evidence, it is possible that SDC1 expression in prostate tumours is part of a recapitulation of developmental signalling programmes, although this would have to be experimentally confirmed.

3.3.3.4 ZSCAN4

Analysis of ZSCAN4 expression variation in a placental dataset revealed an enrichment of GO terms involving metabolism and intracellular transport in samples that highly expressed ZSCAN4. Unfortunately, given the lack of prior information about ZSCAN4 expression in placenta, it is not yet possible to place this finding in a meaningful

biological context. Currently known functions of ZSCAN4 have been characterised in mouse embryonic stem cells, where it is thought to be expressed in self-renewing pluripotent cells (Amano *et al.*, 2013; Sharova *et al.*, 2016). It serves to regulate telomere length and telomere elongation in these cells (Zalzman *et al.*, 2010; Zhang *et al.*, 2016; Nakai-Futatsugi and Niwa, 2016), and appears to be modulated by PI3K signalling (Storm *et al.*, 2014), DNA damage (Storm *et al.*, 2014) and retinoic acid (Tagliaferri *et al.*, 2016; Sharova *et al.*, 2016).

However, such data may be misleading, as they are all characterised in mouse cells. The results obtained from extremes of expression analysis may suggest that ZSCAN4 may have additional roles in metabolism in human datasets, and it would be interesting to see if future functional studies match up with these findings. This data also identifies transcription of ZSCAN4 mRNA in human placenta, and analysis of ENCODE and Illumina BodyMap 2.0 high throughput sequencing data from RNA-Seq, CHIP-Seq and CAGE studies reveals that there may be a longer ZSCAN4 transcript in this tissue – and possibly others – than was previously identified. However, there is a pseudogene just downstream of this putative transcription start site which is too small to represent this transcript in full. Given these findings, experimental characterisation is required to understand the functions of ZSCAN4 in human cells, and it is still very poorly characterised with respect to cancer, with only two papers reporting ZSCAN4 expression in cancer cell lines, where it seems to interact with telomere-related proteins (Lee and Gollahon, 2014, 2015).

3.3.4 Development of a mouse tissue microarray

In order to test the process of constructing tissue microarrays and conducting staining, an FFPE mouse tissue microarray was generated. Given that tissue microarray construction was successful, H&E staining and IHC staining was undertaken.

3.3.4.1 ALDH7A1 staining on the mouse microarray

Anti-ALDH7A1 IHC conducted on the mouse tissue microarray revealed that most tissue types demonstrated consistent nuclear staining with the antibody, except prostate tissue which demonstrated cytoplasmic staining. The change in protein localisation in the prostate suggests that ALDH7A1 is fulfilling a different function or being regulated differently here than in other tissues. ALDH7A1 has both a nuclear localisation signal and a nuclear export signal, as well as an additional mitochondrial transport signal present in a splice isoform (Brocker *et al.*, 2010), and previous functional studies of

ALDH7A1 in cell culture have suggested that changes in localisation from cytoplasm to nucleus occur throughout the cell cycle (Chan *et al.*, 2011). However, Chan *et al.* also show changes in protein localisation from the nucleus to the cytoplasm during phases of the cell cycle, which we did not observe in our tissues: staining was either solely nuclear or solely cytoplasmic. There was considerable variation in nuclear signal intensity, with the most heterogeneous pattern observed in pancreas: negative cells were present next to neighbouring weakly and strongly staining cells. Aldehyde dehydrogenase activity has previously been reported to fluctuate in cell culture conditions, where cells displaying low activity can give rise to cells with high activity and vice versa (Doherty *et al.*, 2011). Although this phenotypic plasticity has not been demonstrated specifically for ALDH7A1, varied expression in tissues could be explained by asynchronous cell cycle progression by cells in the tissue, as ALDH7A1 staining by immunohistochemistry appears to be affected by phases of the cell cycle (Chan *et al.*, 2011). Overall, the ALDH7A1 staining patterns observed by IHC are consistent with the published literature on ALDH7A1 localisation and expression patterns within mouse tissues by western blot and by IHC (Brocker *et al.*, 2010). The antibody reaction conditions are suitably optimised for further testing in human tissue.

3.3.4.2 ZSCAN4 staining on the mouse microarray

Anti-ZSCAN4 IHC conducted on the mouse tissue microarray showed a diffuse and weak cytoplasmic staining in most tissues examined. For example, both liver and pancreas presented diffuse cytoplasmic staining patterns with no preference for cell type. However, portions of the pancreas displayed no staining. This is not consistent with previous descriptions of ZSCAN4 expression in the mouse pancreas, which displayed a small proportion of strongly nuclear and cytoplasmic positive cells in pancreatic tissue including islets (Ko *et al.*, 2013). This is surprising, given that the group used the same antibody clone for their anti-ZSCAN4 immunohistochemistry, present at the same working dilution. However, there is no other literature to corroborate the evidence provided by Ko *et al.* (2013). If expressed at all, ZSCAN4 staining would be predicted to be present in the nucleus given its status as a transcription factor, as well as its documented roles in responding to DNA damage (Storm *et al.*, 2014) and maintaining telomere length (Zalzman *et al.*, 2010; Storm *et al.*, 2014). Therefore, the diffuse staining in all mouse tissues is suspected to be background staining from non-specific primary antibody binding. Taken together, the results suggest that the anti-ZSCAN4 antibody does not suitably recognise the antigen in mouse tissues,

and further testing in human FFPE tissues is the next step for further benchmarking this antibody.

3.3.4.3 Overall utility of the microarray

When subjected to IHC procedures, core loss rate across the entire array averaged at approximately 30%. In some tissue microarray studies, the rate of core loss can be as low as 2% (Tennstedt *et al.*, 2012), but in tissue microarrays constructed from scant tissue such as prostate biopsies, the core loss can be as high as the 24% reported by (Singh *et al.*, 2007). The high rate of core loss in the mouse tissue microarray could be attributed to a lack of tissue depth in many mouse tissue blocks, such as ovary and testis. Another possibility for tissue microarray core loss is removal during the harsh antigen retrieval procedures, although this was not observed to occur in these IHC experiments.

In attempt to resolve the problem of scant tissue, in some cases multiple partial cores were packed into a single core to increase the amount of tissue available for analysis. Although this should not be a problem for radical prostatectomy and TURP chip samples given their large size and specimen thickness, it suggests that prostate biopsies may be insufficient material for tissue microarray construction – a conclusion also supported by Singh et al. (2007), especially when these prostate biopsies have already been sectioned through for prior pathology use and may have a considerably reduced specimen content. In these cases, whole sectioning of single samples would need to be carried out for IHC. However, overall tissue microarray construction was deemed to be successful, with the inclusion of a wide variety of mouse tissues that could be simultaneously subjected to IHC. It is hoped that in the future, in addition to being a useful platform for optimising some antibody staining procedures via IHC, it will also be a useful resource to this and other laboratories for the expression profiling of markers of interest in mouse tissues.

4 Assessment of Candidate Biomarkers in the Bath Cohort

4.1 Introduction

Having identified a group of proteins that could be potential biomarkers for testing by IHC, and optimising the conditions required to obtain a satisfactory IHC signal, potential biomarker expression patterns were evaluated in human prostate tissue from benign and cancerous regions of clinical prostate biopsies and surgical samples. Clinical data was used to determine the relationship between biomarker staining patterns and the clinical features of patients with prostate adenocarcinoma. The main aim was to evaluate the prognostic utility of potential biomarkers to be tested on a small scale, so that candidate biomarkers can be taken forward for further investigation. Hypotheses concerning each potential biomarker are described below.

4.1.1 Hypotheses concerning prospective biomarkers

4.1.1.1 BMI1 expression will be positively associated with biochemical recurrence

BMI1, a Polycomb group protein, is associated with stem cell and self-renewal activity. It is expressed in fractions of prostate cancer cultures that have greater tumour-initiating (Lukacs *et al.*, 2010b) and sphere-initiating capacity (Rybak *et al.*, 2011), and BMI1 expressing cells have been associated with stem-like phenotypes in multiple studies (Zhang *et al.*, 2012a; Hurt *et al.*, 2008; Lukacs *et al.*, 2010b). Overexpressing BMI1 in

prostate cancer cell lines and xenografts results in increased cell proliferation and a resistance to chemotherapeutic drugs, as well as an increase in anti-apoptotic protein BCL2 and Cyclin-D1 (Siddique *et al.*, 2013b). BMI1 is also thought to inhibit prostate cancer cell senescence by repressing transcription of p16Ink4a (Ammirante *et al.*, 2013). Serum levels of BMI1 are positively associated with increased tumour stage (Siddique *et al.*, 2013b; Siddique *et al.*, 2013c). and mRNA levels of BMI1 and other proteins in its pathway have been used as part of a multi-cancer poor prognosis gene expression signature (Glinsky *et al.*, 2005). BMI1 is upregulated following castration in mouse models of prostate cancer and promotes castration-independent and tumour growth, a phenomenon that appears to be dependent on activation of the IkB kinase Ikkα and its subsequent E2F1-mediated transcription of BMI1 (Ammirante *et al.*, 2013). The current literature therefore suggests that BMI1 may be important in the patient response to chemotherapy and hormone therapy. I hypothesise that BMI1 staining levels will be higher in patients that will go on to develop recurrent disease, compared to those with non-recurrent disease.

4.1.1.2 ALDH7A1 will be positively associated with biochemical recurrence

Aldehyde dehydrogenase (ALDH) activity is routinely used as a marker of stem-like cancer cells in multiple types of tumour, including in breast (Wang *et al.*, 2011) and prostate cancer (Vesuna *et al.*, 2009; van den Hoogen *et al.*, 2010). Some aldehyde dehydrogenases catalyse the biosynthesis of retinoic acid, which is an important molecule involved in differentiation (Alison *et al.*, 2010). They also participate in aldehyde and alcohol metabolism. Thus, ALDH enzymes may be functionally involved in self-renewal and resistance to alkylating agents such as cyclophosphamide. Enzymes of the ALDH family are also associated with poor prognosis in a number of cancers (Alison *et al.*, 2010). In particular, ALDH1A1 and other stem cell markers are upregulated in castration resistant prostate cancer compared to non-castrated metastatic disease (Pfeiffer *et al.*, 2011), suggesting a role for ALDH1A1 in the response to androgen deprivation. Indeed, high ALDH1A1 expression in prostate cancer specimens has been associated with cancer stem-like phenotypes, higher Gleason scores and stages, as well as a poor prognosis (Li *et al.*, 2010). It is therefore possible that other ALDH enzymes are implicated in prostate CSCs and may have utility in prostate cancer prognostication. In prostate cancer cell lines, knockdown of ALDH7A1 expression has been found to reduce expression of regulators of epithelial-mesencyhmal transition and to reduce the putative cancer stem cell subpopulation (van den Hoogen *et al.*, 2011),

with a concomitant decrease in bone metastatic capacity. Currently there are no studies investigating its relationship with recurrence or patient outcome. However, due to its influence on invasion and aggressiveness of tumours, I hypothesise that ALDH7A1 will be expressed more strongly in patients that go on to experience recurrence than in non-recurrent patients.

4.1.1.3 SLC31A1 will be positively associated with biochemical recurrence

SLC31A1 was identified in this study as being overexpressed at the mRNA level in patients with recurrent prostate tumours in Chapter 3.3.2. SLC31A1 is the major transporter for copper uptake in humans (Jong and McKeage, 2014). Maintaining copper homeostasis is important in both benign and malignant prostate cancer cells, as copper chelation reduces cell viability in both cases (Safi *et al.*, 2014). Both prostate cancer cell lines and xenografts all having a high accumulation of copper and, though the reason for this accumulation is unknown, it allows for effective pharmacological targeting of castration-resistant prostate xenografts and cell lines (Safi *et al.*, 2014). AR activity positively regulates the expression of SLC31A1 and subsequent copper intake by the VCaP cell line (Safi *et al.*, 2014). Additionally, there may be a link between SLC31A1 and ERK signalling, as mouse embryonic fibroblasts with a homozygous loss of SLC31A1 fail to phosphorylate and activate ERK when stimulated by ligands of its upstream receptor tyrosine kinases, such as FGF, PDGF and EGF (Tsai *et al.*, 2012). This suggests that intracellular copper is important for activating the ERK pathway that is commonly dysregulated in cancer. However, a survey of normal and malignant human tissue for SLC31A1 protein levels by IHC found no detectable levels of the protein in either benign prostate or prostate carcinoma (Holzer *et al.*, 2006). The data on SLC31A1 in prostate cancer is limited and, given the findings in Chapter 3 and the description of SLC31A1 expression in prostate cancer cell lines, further investigation is warranted. I predict that SLC31A1 will be more highly expressed in patients with recurrent tumours than those that remain in remission.

4.1.1.4 The C-terminal domain of MUC1-C will be positively associated with biochemical recurrence

Mucin 1 transmembrane protein (MUC1) was initially identified as being relevant to prostate cancer in a microarray experiment attempting to subset prostate cancer patients into clinically relevant categories (Lapointe *et al.*, 2004). The functions of MUC1 outside of its normal roles in mucous membrane formation are incompletely understood, as MUC1-null mice do not have a discernible phenotype in a pathogen-free

environment (Hattrup and Gendler, 2008). However, MUC1 is cleaved during protein folding and the C-terminal domain (MUC1-C) is thought to serve as a signalling molecule which binds to kinases such as EGFR family members, as well as non-kinases such as beta-catenin (Hattrup and Gendler, 2008). Phosphorylation of tyrosines on MUC1-C can enhance its binding to beta-catenin, resulting in its loss from adherens junctions, enhanced beta-catenin nuclear activity and thus facilitating anchorage-independent growth (Hattrup and Gendler, 2008).

With respect to prostate cancer, MUC1 expression was associated with a more aggressive subset of the disease in a clustering analysis (Lapointe *et al.*, 2004). Subsequent IHC on a tissue microarray of 225 primary prostate tumours revealed that increased MUC1 protein levels are associated with a shorter time to recurrence and are an independent predictor of cancer recurrence (Lapointe *et al.*, 2004). It was later found to be expressed in the CSC-like side population of MCF7 breast cancer cells which showed other putative cancer stem cell characteristics such as being CD44+/CD24- (Engelmann *et al.*, 2008). Castration-resistant cell lines DU145 and PC-3 express MUC1, whereas Androgen-sensitive LNCaP cells do not (Joshi *et al.*, 2009). In MUC1-expressing prostate cancer cell lines DU145 and PC3, inhibition of oligomerisation of MUC1-C results in necrotic cell death and reduces cell proliferation, while in xenografts this treatment strikingly induces complete regression of tumours (Joshi *et al.*, 2009), suggesting that the molecular activity of MUC1-C is of potential clinical relevance. Furthermore, MUC1-C expression was found to suppress AR expression, as well as associating with AR and binding to the PSA promoter (Rajabi *et al.*, 2012); overexpression of MUC1-C in LNCaP cells resulted in increased invasion, upregulation of vimentin and downregulation of E-cadherin that would be associated with an epithelial-mesenchymal transition, as well as androgen-independent growth (Rajabi *et al.*, 2012). The current evidence suggests that MUC1-C is at least partly responsible for the association between MUC1 and aggressive prostate tumours. Given this information, I hypothesise that MUC1-C will be more highly expressed in patients that have experienced biochemical recurrence within 5 years than those that remain in remission.

4.1.1.5 Nestin will be positively associated with biochemical recurrence

The utility of the intermediate filament Nestin as a biomarker has been demonstrated before, in a dual IHC setting to identify microvascular proliferation in prostate tumours

alongside Ki-67 (Gravdal *et al.*, 2009). High expression of Nestin in this context is associated with biochemical failure, recurrence and metastasis. Nestin expression alone maybe associated with high grade tumours and the presence of metastasis (Jani *et al.*, 2010). However, Nestin expression is generally also associated with the progression of tumours (Kleeberger *et al.*, 2007a) and tumour cell lines (Pfeiffer *et al.*, 2011) to an androgen-independent state. I hypothesise that Nestin will be more highly expressed in patients that go on to experience recurrence than in non-recurrent patients.

4.1.1.6 ZSCAN4 will be expressed in malignant but not benign tissue, and will be positively associated with biochemical recurrence

There is currently no clinical analysis on the expression of the pluripotency-related transcription factor, ZSCAN4, in tumours, to the author's knowledge. However, it is known that ZSCAN4 has a very restricted expression pattern during development, being expressed at the 2-cell stage of mouse embryonic development and in embryonic stem cells (Falco *et al.*, 2007), where it controls telomere elongation (Zalzman *et al.*, 2010), the response to DNA double-strand breaks (Storm *et al.*, 2014) and also long-term self-renewal ability (Storm *et al.*, 2014; Amano *et al.*, 2013). Gollahon *et al*. have since proposed that ZSCAN4 performs similar telomere-related roles in cancer cell lines (Lee and Gollahon, 2014, 2015). In mouse embryonic stem cells, ZSCAN4 is regulated by PI3K signalling, and introduction of a specific inhibitor that targets the p110α catalytic isoform reduces ZSCAN4 expression by 80% (Storm *et al.*, 2014). Given its highly-restricted expression pattern, its relationship with pluripotency and its association with PI3K signalling, we hypothesise that ZSCAN4 will be exclusively expressed in malignant but not benign prostate epithelium, and that its expression will be greater in patients that go on to experience recurrence than in non-recurrent patients.

4.1.1.7 RS1 will be positively associated with biochemical recurrence

The secreted extracellular protein RS1 has not been studied with respect to cancer. Instead, its functions are largely studied in the eye, where it is involved in organising cells of the retina and mutations of this protein lead to X-linked Retinoschisis (Xu *et al.*, 2011). Prior data in this thesis suggested that it is overexpressed at the mRNA level in the prostate tumour tissue of patients that would later experience recurrence, and is amplified in metastatic and neuroendocrine prostate tumours (Chapter 3.2.2). Using this novel data, we predict that this trend will continue at the protein level, and RS1 will be more highly expressed in the primary tumour tissue of patients that have suffered recurrence.

4.1.2 Androgen Receptor expression and its association with prognosis

Androgen Receptor (AR) signalling is a key driver in prostate cancer progression and its path to recurrence (Chapter 1.2.3), with mutations and amplifications of the AR gene frequently being acquired during the acquisition of castration resistant phenotypes, as well as the deregulation of AR-mediated transcriptional control through alterations in co-activators such as ERG. On the other hand, some putative prostate CSCs are AR-negative, which would also be consistent with their enrichment upon androgen withdrawal (Chapter 1.3.4). Reducing or perturbing AR signalling can induce the expression of putative cancer stem cell markers such as MSI-1, Nanog, CD44 and SOX2 (Schroeder *et al.*, 2014). Conversely, the expression of the stem cell marker GLI1 in the androgen-dependent cell line LNCaP causes AR expression to be lost and an androgen-independent phenotype is acquired (Nadendla *et al.*, 2011). In prostate tumours, AR staining assessed by IHC positively correlates with the expression of proliferative markers Ki-67 and phosphorylated histone H3 (Goltz *et al.*, 2015), and the combination of these factors has greater prognostic relevance than the proliferative markers alone. Given that alterations of AR expression and activity are of importance to prostate cancer and to prostate cancer stem cells, measuring AR levels and correlating this expression to the expression of potential biomarkers in this chapter would help to further characterise the regulation of these markers and, potentially, add prognostic utility.

4.1.3 Aims

Now that prospective biomarkers have been chosen and hypotheses assigned, their utility as biomarkers and their expression in prostate cancer must be addressed. This chapter will address the following aims:

i. To determine the expression patterns and levels of potential biomarkers in benign and malignant prostate tissue, and the ability of biomarkers to distinguish between normal and disease states.

ii. To identify if candidate biomarkers are differentially expressed between patients that will relapse within 5 years and those that will remain in remission.

iii. To assess any associations between candidate biomarkers and androgen receptor expression status.

4.2 Results

4.2.1 Study Design

4.2.1.1 Objectives

The first step in testing was to design the study, produce a protocol and submit it for NHS ethical and Research & Development approval. In order to test the relationship between the chosen candidate markers and recurrence, study was designed to compare recurrent patients and patients that have not recurred after 5 years, in each of two treatment modalities: prostatectomy and hormone therapy. The primary endpoint of the study was biochemical (PSA) recurrence at 5 years from the date of diagnosis.

4.2.1.2 Principle Inclusion/Exclusion Criteria

To assess patients records for their suitability in the study, inclusion and exclusion criteria were established and published in a study protocol (Appendix 5) that was approved prior to consenting patients.

To be included in the study, patients met all of these criteria: Older than 18 years of age at the time of diagnosis; Were male; Had a minimum of 5 years of retrospective follow-up time (in the case of non-relapsing patients); Must have had no evidence of metastatic disease at the time of sample acquisition; Had stored clinical data, to include: age, Gleason grading, TNM staging, treatment modalities used, and time to biochemical recurrence; Must have signed written informed consent forms as detailed by the Human Tissue Act of 2004.

Patients were excluded from the study if they met any of these criteria: Patients have had previous or concurrent cancer(s) that is distinct in primary site of histology from the tumour being evaluated, with the exception of treated basal cell carcinomas or superficial bladder tumours; Patients have not been assessed for lymph node involvement or metastases; Patients treated with neoadjuvant hormone therapy or chemotherapy prior to their primary therapy; Insufficient amounts of archived tumour tissue available for the patient to conduct the assays; Patients that have withdrawn informed consent for use of their archived tissue at any time during the study.

The inclusion criteria ensure that the obtained data is suitable for downstream analysis and that protocols are compliant with the Human Tissue Act 2004. Exclusion criteria exclude patients that would present with confounding variables or introduce additional sources of bias. For example, neoadjuvant therapies could alter the expression of biomarkers prior to biopsy and alter patient survival outcomes. Similarly, a co-existing tumour and/or prior cancer therapies have the potential to alter treatment choices, survival profiles and the biomarker expression profiles of the patients.

4.2.1.3 Patient identification and consent

Patient records were searched and suitable candidates identified by RUH clinicians Dr Mark Beresford and Dr Rebecca Bowen. Patients were contacted using a one-page study invitation letter (Appendix 6) inviting them to learn more about the study. They were then presented with patient information sheets (Appendix 7) briefly summarising the purposes of the study, and detailing any study risks, benefits and considerations for participants. Patients were then asked if they wished to sign a consent form (Appendix 8) for inclusion in the study, where their anonymised clinical data and archived tissue samples would be taken for the subsequent analysis. Samples were stored securely on site at the Department of Biology and Biochemistry, University of Bath, under conditions that when the study ends, the remaining samples return to the Department of Pathology at the RUH or be destroyed.

4.2.1.4 Sample size calculations, protocol and changes to objectives during study development

Estimated required sample sizes were determined by power calculations performed by Dr Gordon Taylor (Reader in Medical Statistics, University of Bath) using estimated assay sensitivity and specificity values of 80% for distinguishing recurrent from non-recurrent patients. It was estimated that 20 patients would be required for each test

group to resolve this level of difference. Therefore, initial plans were to collect 20 patients for each of the relapsing and non-relapsing categories, for each of the two treatment groups, for a total of 80 patients. The initial aim was to construct a tissue microarray including single cores of tissue from each patient. However, due to the scant nature of tissue taken from prostate biopsies in this study, whole sections of each tissue specimen were evaluated by IHC rather than constructing a tissue microarray on which to conduct staining.

NHS ethical and Research & Development approval was obtained (REC reference: 13/WS/0153; IRAS project ID: 112241). Tissues were obtained from the primary tumour at or prior to time of treatment (either by biopsy or obtained from prostatectomy) or from transurethral resection of the prostate for symptomatic relief. Because of the small number of samples in the Bath Cohort, whole sections of each tissue specimen were evaluated by IHC rather than constructing a tissue microarray to conduct the staining on.

Identifying suitable patient records was time-intensive, as patients diagnosed from the time period of 1997-2012, when the samples were initially taken, had their clinical information stored in a variety of different clinical databases, which made it difficult to compile. Due to the limited amount of samples that could be obtained within a practicable time frame, the pilot study was reduced in size to a total of 27 patients (10 non-recurrent and 17 recurrent) with a mix of primary treatment modalities representative of prostate cancer patient intake at the RUH. The full structure of the Bath Cohort is described in the next section.

4.2.2 Structure of the Bath Cohort

The Bath patient cohort consisted of 27 patients diagnosed with prostate cancer at the Royal United Hospital, Bath, between 1997 and 2012. The cohort was retrospectively acquired, with samples obtained from a range of surgical or diagnostic procedures, depending on the individual treatment course of the patient. Samples were: diagnostic core needle biopsies (22/27, 81.5%); portions of the radical prostatectomy specimen (3/27, 11.1%); or chips from transurethral resection of the prostate (2/27, 7.4%). Available data included age, TNM staging, Gleason grading, PSA levels at diagnosis, and biochemical recurrence status at 5 years. A summary of these data is presented in Table 4.1. The majority of patients were treated with radiotherapy (10/27, 37.0%) or hormone therapy (6/27, 22.2%) as a primary modality, but other treatment options

included surveillance and radical prostatectomy. No patients in this cohort were assigned a primary Gleason grade of 5. Two patients (7.4%) presented with metastatic disease at the time of diagnosis. Subsequent therapies varied widely and overlapped in their usage, with the three most common treatment option being hormone therapy (18/27, 66.7%), Radiotherapy for palliative or non-palliative intent (6/27, 22.2%), and Docetaxel (3/27, 11.1%). A summary of these data is presented in Table 4.2.

Table 4.1: Clinical characteristics of patients from the Bath cohort.

	Frequency/Statistic	***%***
Number of Patients	27	
Age:		
Range	58-81	
Mean	67.7	
Median	67.3	
<50	0	4.2
50-59	1	3.7
60-69	14	51.9
>=70	9	33.3
Primary Gleason Grade:		
3	14	51.9
4	11	40.7
5	0	0
ND	2	7.4
Overall Gleason Grade:		
Mean	7.04	
<=Gleason 3+4	14	51.9
>=Gleason 4+3	11	40.7
ND	2	7.4
T Category:		
T1	7	25.9
T2	11	40.7
T3	6	22.2
T4	1	3.7
ND	2	7.4
N Category:		
N0	21	71.8
N1	1	3.7
N2	1	3.7
ND	4	14.8
M Category:		
M0	17	63.0
M1	2	7.4
ND	8	29.6
Biochemical Recurrence Status (at 5 years):		
Non-Recurrent	10	37.0
Recurrent	17	63.0
Specimen Type:		
Radical Prostatectomy	3	11.1
TURP Chips	2	7.4
Biopsy Cores	22	81.5

Table 4.2: Treatment characteristics of patients from the Bath Cohort

	Frequency	*%*
Number of Patients	27	
Primary Therapies:		
Radiotherapy	10	37.0
Radical prostatectomy	3	11.1
Hormone therapy	6	22.2
Chemotherapy	1	3.7
Surveillance	4	14.8
Data not available	3	11.1
Subsequent Therapies:		
Hormone Therapy (LHRHa, MAB, Antiandrogens)	18	66.7
Radiotherapy (non-palliative)	4	14.8
Radiotherapy (palliative)	2	7.4
High Intensity Focused Ultrasound	1	3.7
Docetaxel	3	11.1
Carboplatin + Etoposide	1	3.7
Samarium	1	3.7
PSA Nadir (if treated during follow-up):		
<=0.5ng/mL	16	59.3
0.6-1.0ng/mL	1	3.7
>1.0ng/mL	5	18.5
N/A	2	7.4
Data not available	3	11.1

4.2.3 Expression Patterns of Candidate Biomarkers in Human Prostate Tissue

4.2.3.1 ALDH7A1 is expressed in the cytoplasm and nucleus of both benign and malignant prostate glands

Expression of ALDH7A1 was first investigated by immunohistochemistry on prostate biopsies and radical prostatectomy blocks, and staining in histologically benign tissue was compared against carcinoma tissue. ALDH7A1 was expressed in both benign (Figure 4.1A and 4.1B) and malignant prostate epithelia with widely varying staining intensities in both groupings. Staining intensity ranged from strong (Figure 4.1C) to intermediate (Figure 4.1D), weak (Figure 4.1E) and negative (Figure 4.1F). As reported before in Immunohistochemical studies in mouse (Brocker *et al.*, 2010), ALDH7A1 is expressed in the kidney and the human tissue also appears to serve as a positive control with cytoplasmic staining in the proximal renal tubules (Figure 4.1G). Negative controls, where no primary antibody was added, never resulted in a positive staining reaction (Figure 4.1H).

Localisation was predominantly cytoplasmic in all cases, but nuclear staining of ALDH7A1 was also present at a low and variable frequency in both benign and malignant prostate epithelium (Figure 4.1, arrows). Cytoplasmic staining was generally uniform across prostate (Figure 4.1A) and carcinoma glands (Figure 4.1C), with expression in both the basal and luminal cell layers of benign prostate epithelium and no staining in stromal areas. There were no significant differences observed in the mean scored intensity of the cytoplasmic stain ($p=0.6844$) or in the frequency of epithelial cells staining nuclear positive between benign and malignant samples ($p=0.5835$) (Figure 4.2A and B respectively). There was no significant difference in ALDH7A1 nuclear or cytoplasmic expression between low Gleason grade ($\leq 3+4$) or high Gleason grade ($\geq 4+3$) tumours ($p=0.5969$ and 0.8691 respectively).

4.2.3.2 ALDH7A1 staining does not predict 5-year biochemical recurrence status

In order to assess the potential influences of ALDH7A1 immunoreactivity at biopsy on patient outcome, patients were divided into two groups based on the time from diagnosis to biochemical recurrence. Neither mean ALDH7A1 nuclear H-Score ($p=0.6661$) nor mean cytoplasmic score ($p=0.8763$) were significantly different between 5-year recurrent and non-recurrent patients, and there was comparable heterogeneity in ALDH7A1 staining among patients in both groups (Figure 4.2C and D). Therefore, the data is not consistent with a significant effect of ALDH7A1 expression on recurrence.

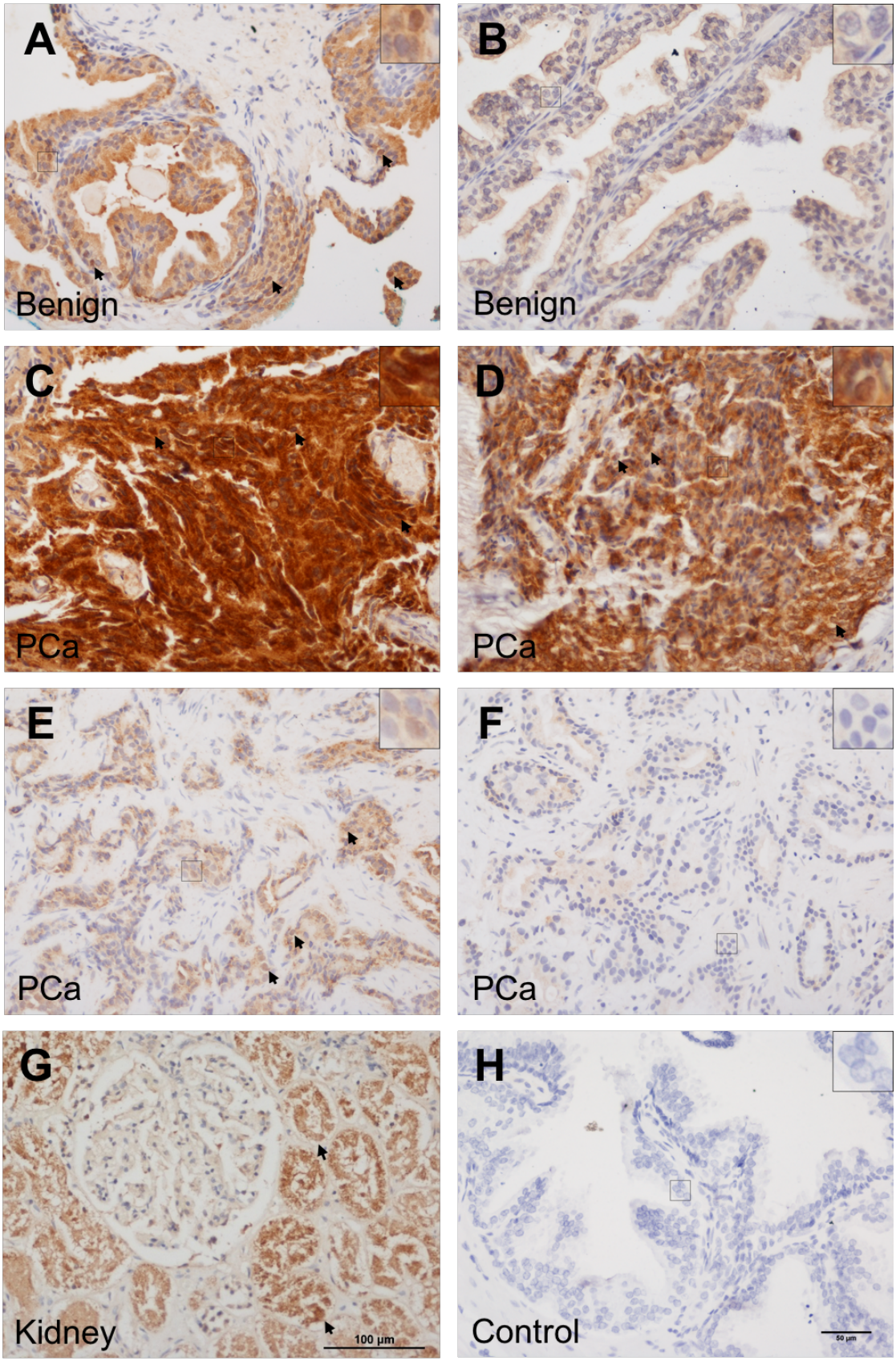


Figure 4.1: Anti-ALDH7A1 IHC in benign and malignant prostate tissue.

ALDH7A1 is expressed variably in the nuclei and cytoplasm of benign prostate epithelium, ranging from a moderate cytoplasmic intensity and some nuclear staining (A, arrows) to weak cytoplasmic staining intensity and no nuclear staining (B). ALDH7A1 expression in prostate cancer (PCa) also occurs in the nuclei and cytoplasm of the epithelial cells, but ranges in intensity from strongly staining (C) with many positive nuclei (arrows), to intermediate staining (D, arrows), weak cytoplasmic staining with occasional nuclear expression (E, arrows) and negative staining of both nuclear and cytoplasmic compartments (F). ALDH7A1 is expressed in renal tubular epithelium as a positive control (G), whilst a section of prostate tissue with no primary antibody added was used as a negative control and showed no staining (H). Antibody binding is visualised with DAB (brown) and nuclei are counterstained with haematoxylin (blue). Representative fields of view are shown at 20x magnification, with representative epithelial cells shown inset. Scale bars are labelled with their respective sizes.

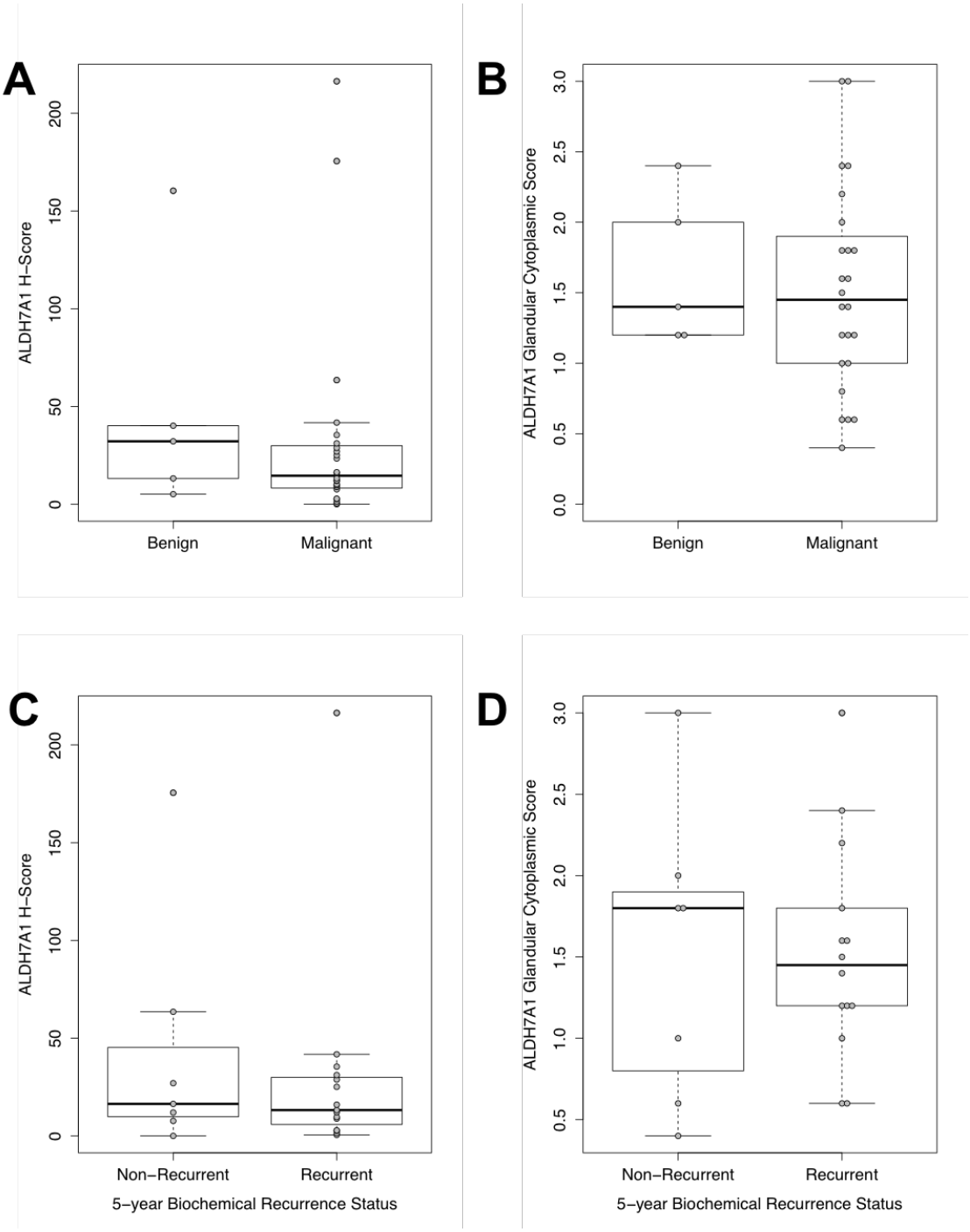

A
ALDH7A1 H−Score
0
50
100
150
200
Benign
Malignant
B
ALDH7A1 Glandular Cytoplasmic Score
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Benign
Malignant
C
ALDH7A1 H−Score
0
50
100
150
200
Non−Recurrent
Recurrent
5−year Biochemical Recurrence Status
D
ALDH7A1 Glandular Cytoplasmic Score
0.5
1.0
1.5
2.0
2.5
3.0
Non−Recurrent
Recurrent
5−year Biochemical Recurrence Status

Figure 4.2: Quantification of nuclear and cytoplasmic ALDH7A1 staining in benign prostate and prostate adenocarcinoma tissue.

IHC staining of ALDH7A1 was quantified in the Bath cohort using the H-Score for nuclear stains and the cytoplasmic score. Boxplots of ALDH7A1 (A) nuclear H-score and (B) cytoplasmic score were compared between benign (n=5) and malignant prostate tissue (n=24), showing no significant difference between the two groupings. 5-year recurrence-free (n=7) and recurrent (n=15) patients were also compared for ALDH7A1 (C) nuclear H-Score and (D) cytoplasmic score, where there was no significant difference. Data represent the mean of five random fields of view per sample. Two-sample t-tests were conducted to determine statistical significance for each set of conditions.

4.2.3.4 Anti-SLC31A1 antibody shows intense background staining

Four prostate biopsies from the newly available Bath Cohort were stained with an antibody recognising human SLC31A1 in order to optimise it, as it would be unlikely to give a good staining result in mouse tissue due to its species specificity. SLC31A1 staining was intense in the glandular epithelium of prostate tumour tissue (Figure 4.3A), but also in the surrounding stromal area (Figure 4.3B), making interpretation of the stain difficult, even though no primary antibody controls turned out negative (Figure 4.3C). This could be explained by a ubiquitous expression of the transporter, but an important feature of any candidate biomarker is a clearly interpretable staining result. Because of this, the marker was excluded from further analysis.

4.2.3.5 MUC1 C-terminal antibody shows no staining in prostate specimens

To ensure that the anti-MUC1 C-terminal antibody was reliable for usage in a large number of precious prostate specimens, its staining was tested with four prostate tumour biopsies using staining protocols optimised by the antibody manufacturer. However, this antibody displayed negative (Figure 4.4A) or weak diffuse cytoplasmic staining (Figure 4.4B) of glandular prostate tumour epithelium that was deemed insufficient for downstream analysis. No primary antibody negative controls, as before, were clear of background staining (Figure 4.4C). In light of these data, MUC1-C was excluded from further analysis.

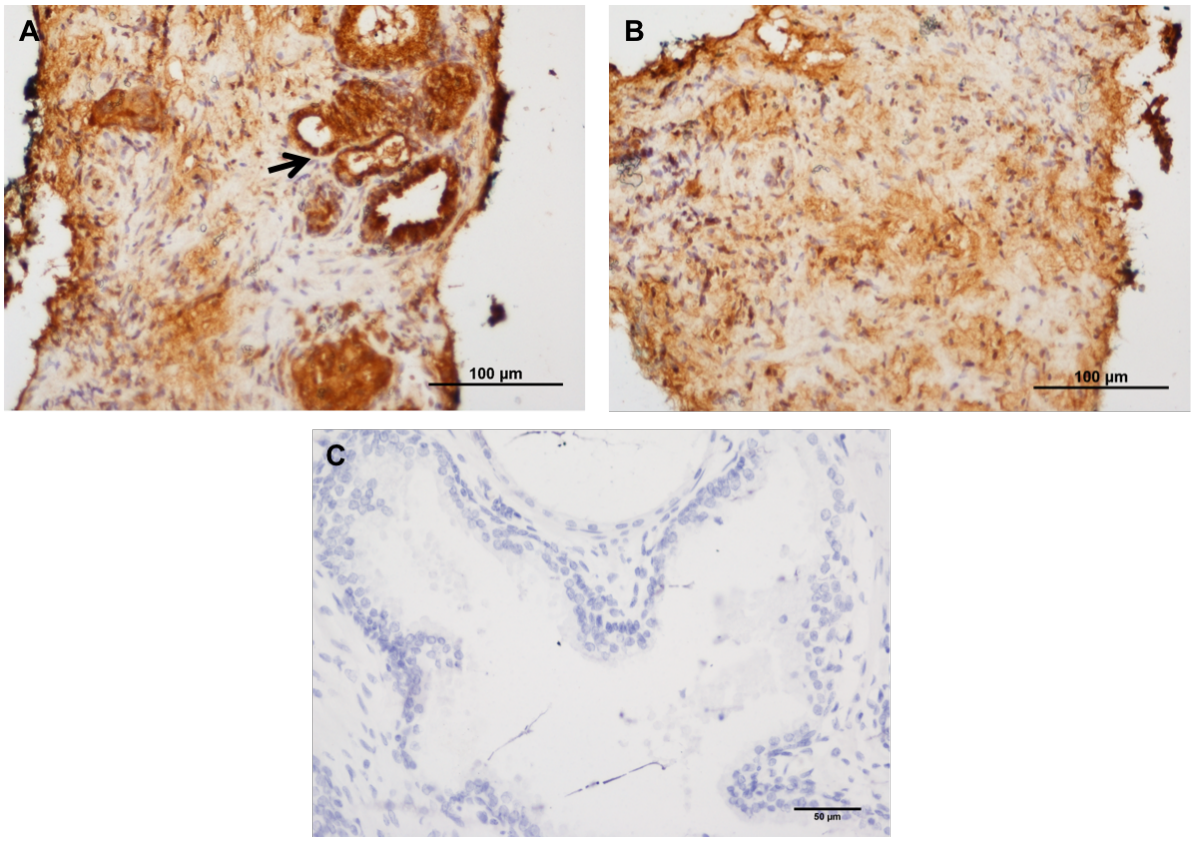


Figure 4.3: SLC31A1 expression in human prostate tumour tissue.

(A) Prostate tumour tissue shows strong cytoplasmic staining of glandular epithelium (arrow) in addition to (B) widespread, diffuse moderate staining in surrounding stromal tissue. (C) Negative control is prostate tumour tissue with no primary antibodies added. All staining performed in a single experiment due to limited sample material. Scale bars – 100µm (A & B) or 50µm (C).

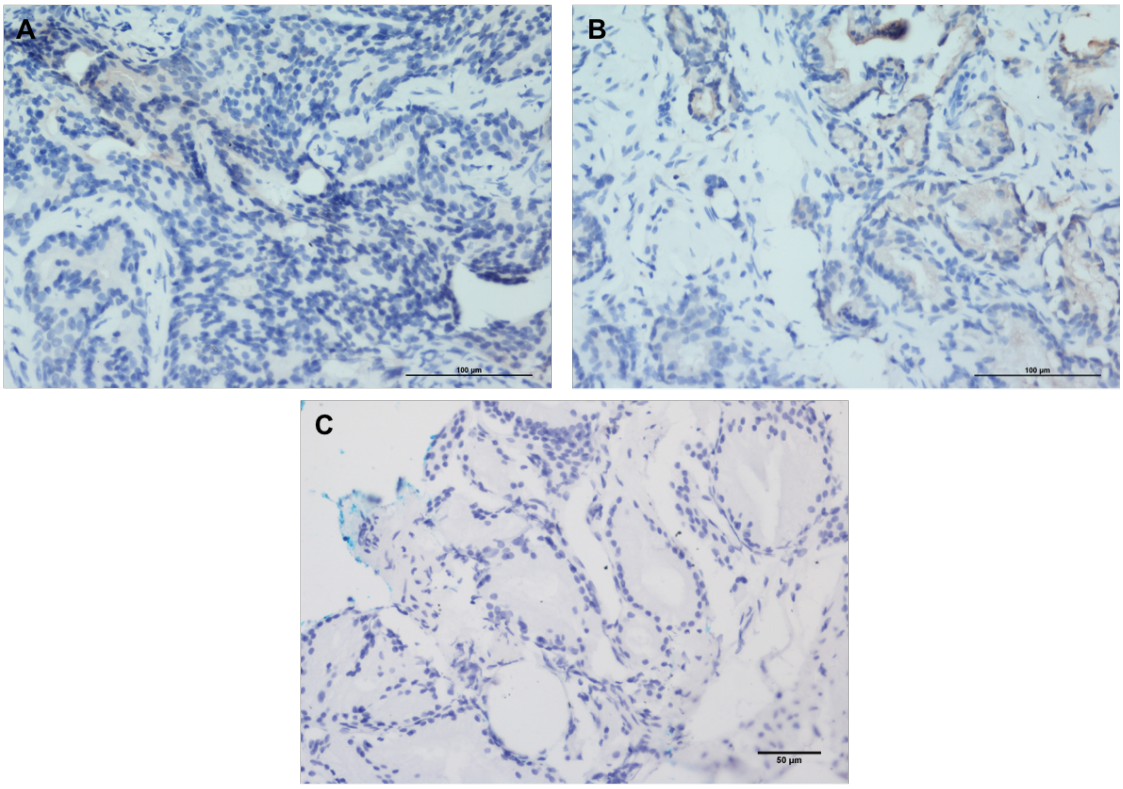


Figure 4.4: MUC1 C-terminal domain expression in human prostate tumour tissue.

(A) Prostate tumour tissue shows no staining in glandular epithelium using an antibody recognising the C-terminal domain of MUC1. (B) Weak cytoplasmic staining of glandular epithelia in another prostate tumour specimen. (C) Negative control is prostate tumour tissue with no primary antibodies added. All staining is performed in a single experiment due to limited sample material. Scale bars – 100µm (A & B) or 50µm (C).

4.2.3.6 Nestin is weakly expressed in both benign and malignant prostate tissue

Anti-Nestin immunohistochemistry was conducted on benign and malignant prostate samples and scored by the cytoplasmic scoring method for intensity alone. A diffuse and uniform cytoplasmic staining reaction was observed in the glandular epithelium of both histologically benign and malignant samples. Cytoplasmic staining was either negative (Figure 4.5A) or weak (Figure 4.5B) in benign epithelium, with infrequent staining of stromal cells in close proximity to epithelium (Figure 4.5A, arrows). Prostate tumours displayed a range of staining intensities from intermediate (Figure 4.5C) to weak (Figure 4.5D) to negative (Figure 4.5E). Diffuse stromal staining was also observed in some malignant prostate samples, which was observed to be of similar intensity to the corresponding epithelial stain and so was not scored (Figure 4.5C and D). As expected, Nestin was detected in glomeruli within the kidney (Figure 4.5F) and in nerves within the prostate tumours (Figure 4.5G), serving as positive controls for the reliability of the immunostain. Negative (no primary antibody) controls verified that the background from the detection system was very low (Figure 4.5H). There was no significant difference in scored cytoplasmic staining intensity between benign and malignant prostate tissue ($p=0.7988$) (Figure 4.6A). There was also no association between Nestin expression and whether tumours were of a low Gleason grade ($\leq 3+4$) or a high Gleason grade ($\geq 4+3$) ($p=0.6528$).

4.2.3.7 Nestin is not significantly associated with 5-year biochemical recurrence status

Nestin immunoreactivity scores were compared between non-recurrent and recurrent patient groups using 5 years or more of biochemical recurrence-free survival as the cut-off (Figure 4.6B). No significant difference was observed in Nestin expression between these two patient groups ($p=0.7383$), indicating no discernible association between Nestin epithelial staining and the response of patients over the first five years of treatment.

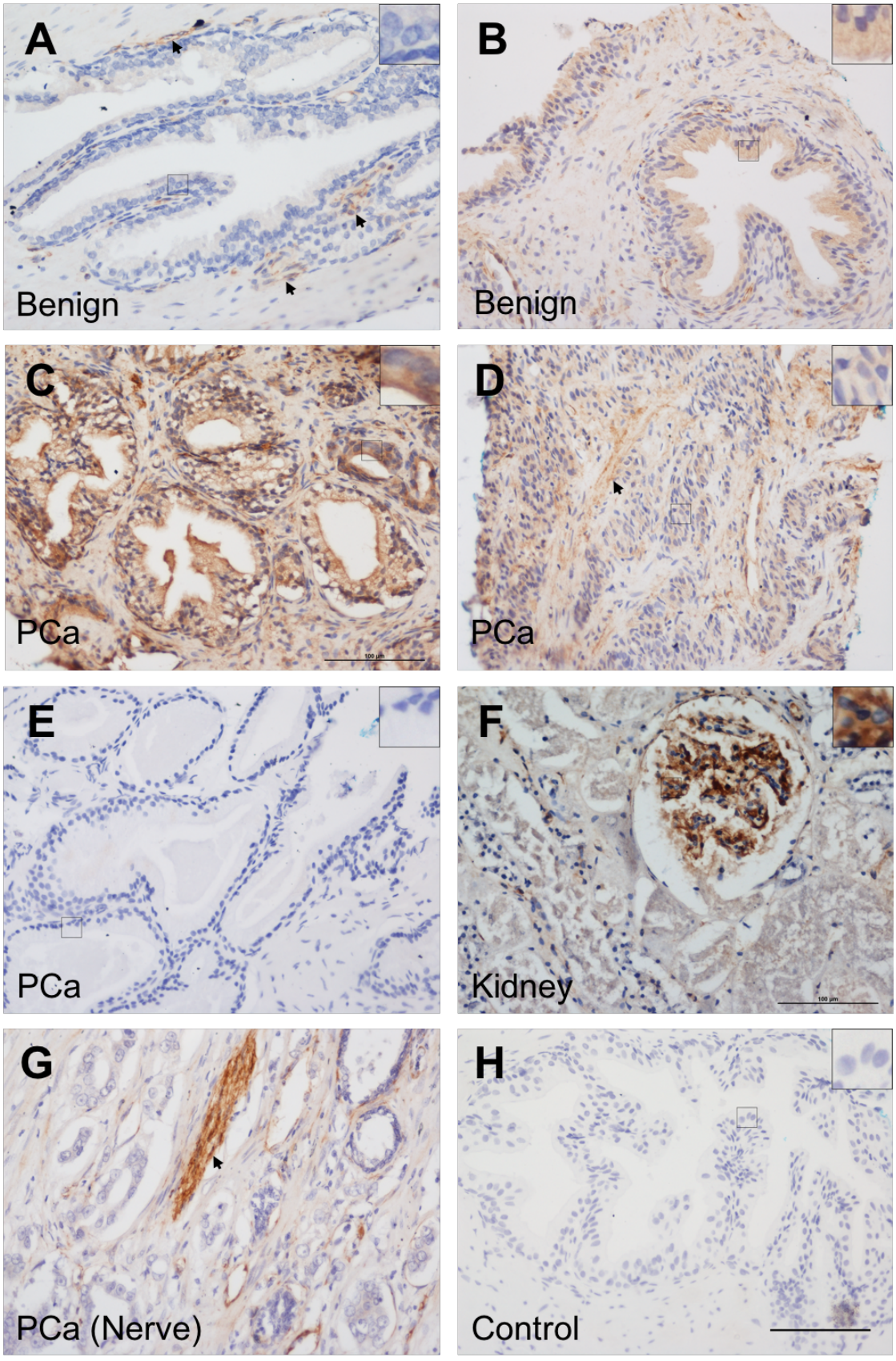


Figure 4.5: Anti-Nestin IHC in benign and malignant prostate tissue

Nestin staining is either negative (A) or weak (B) in the cytoplasm of benign prostate epithelium, although weak staining was observed in stromal cells in some cases (A, arrows). Malignant tissue (PCa) displays a similar cytoplasmic staining pattern ranging from intermediate (C) to weak (D) and negative (E) cytoplasmic staining in the tumour epithelium with some diffuse stromal staining. Glomeruli in the kidney (F) and nerves within the prostate tumours (G) were Nestin positive as expected, demonstrating the specificity of the antibody. Negative controls are free of background staining, as shown by the no primary antibody control in prostate carcinoma tissue (H). Antibody binding is visualised with DAB (brown) and nuclei are counterstained with haematoxylin (blue). Representative fields of view are shown at 20x magnification, with representative epithelial cells shown inset. Scale bar – 100μm.

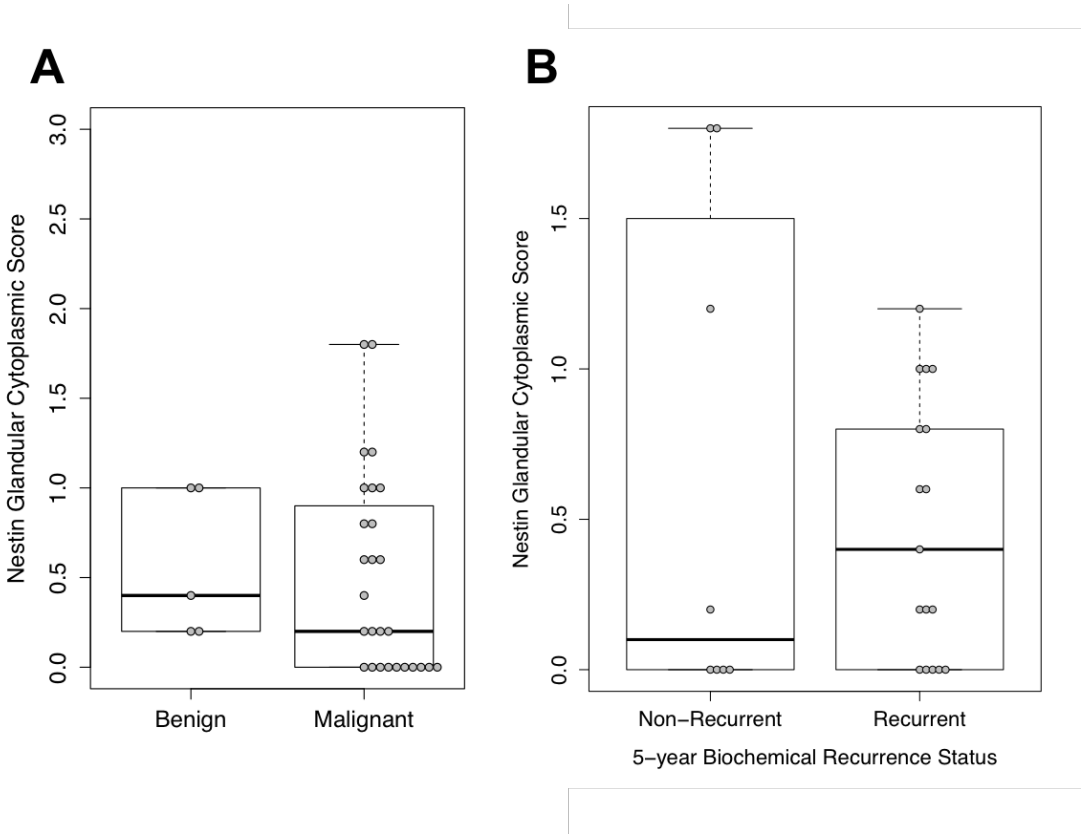


Figure 4.6: Quantification of Nestin cytoplasmic staining in benign and malignant prostate tissue.

IHC staining of Nestin was quantified in the Bath cohort using the Cytoplasmic Score. Boxplots of Nestin Cytoplasmic Score were compared between benign (n=5) and malignant prostate tissue (n=27) (A), showing no significant difference between the two groupings. 5-year recurrence-free (n=8) and recurrent (n=15) patients were also compared for Nestin Cytoplasmic Score (B), where there was no statistically significant difference. Data represent the mean of five random fields of view per sample. Two-sample t-tests were conducted to determine statistical significance for each set of conditions.

4.2.3.8 BMI1 is expressed in the nuclei of benign and malignant prostate epithelia
Histologically benign and malignant prostate tissue was subjected to anti-BMI1 immunohistochemistry and scored by the nuclear H-score method. Consistent with previously published findings (van Leenders *et al.*, 2007), BMI1 expression was observed in the nuclei of benign prostate tissue (Figure 4.7A), but was widely variable between patients with some having negative or low BMI1 H-scores (Figure 4.7B). Malignant prostate tissue also had widely varying frequencies of nuclear BMI1 staining, from strong widespread staining (Figure 4.7C and D, arrows) to weak and scattered staining (Figure 4.7E, arrows) and negative staining (Figure 4.7F, arrows). Negative control prostate tissue with no added primary antibodies came out as consistently negative (Figure 4.7G). There was no significant difference between the mean BMI1 nuclear H-scores of benign and malignant prostate tissue samples, which were 20.3 and 16.4 respectively ($p=0.7178$) (Figure 4.8A). Interestingly, paired benign and malignant samples from the same patient both showed high nuclear BMI1 H-Scores (Figure 4.7A and C respectively). However, there was no significant difference in BMI1 expression between low Gleason grade ($\leq 3+4$) or high Gleason grade ($\geq 4+3$) tumours ($p=0.7369$).

4.2.3.9 BMI1 has lower mean expression in 5-year biochemical recurrent samples
To assess the relationship of BMI1 with relapse in the Bath cohort, patients were again dichotomised on 5-year biochemical recurrence status. There was a trend towards lower BMI1 expression in patients that experienced biochemical recurrence before five years compared to those that had not (Figure 4.8B), with mean H-Scores of 18.7 and 12.1 respectively, although this difference was not statistically significant ($p=0.337$). This observed trend, while not significant, warrants investigation with a larger cohort of patients.

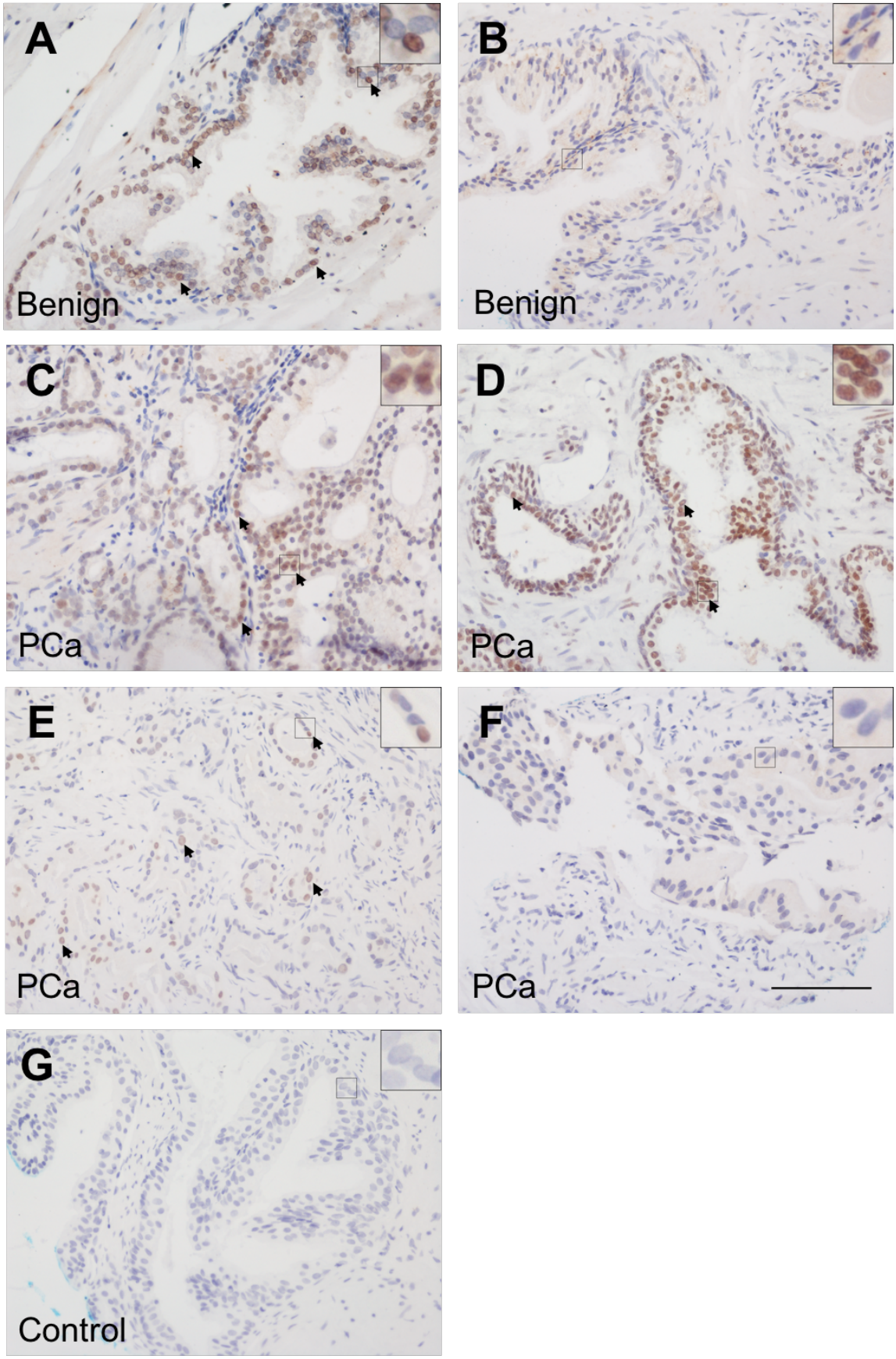


Figure 4.7: Anti-BMI1 IHC in benign and malignant prostate tissue.

BMI1 is expressed heterogeneously in the nuclei of benign prostate epithelium, ranging from scattered, moderately to strongly-staining nuclei (A, arrows) to no nuclear staining (B). Similarly, prostate carcinoma (PCa) samples showed large variation of BMI1 staining. The sample in (C) is the paired tumour region from the sample in (A), showing a similar expression pattern to the matched benign sample. Staining was also heterogeneous in carcinoma epithelium, showing a range from strong nuclear staining (D) to weak and scattered staining (E) and a complete absence of nuclear stain (F). Negative controls are free of background staining, as shown by the no primary antibody control in prostate carcinoma tissue (G). Antibody binding is visualised with DAB (brown) and nuclei are counterstained with haematoxylin (blue). Representative fields of view are shown at 20x magnification, with representative epithelial cells shown inset. Scale bar – 100μm.

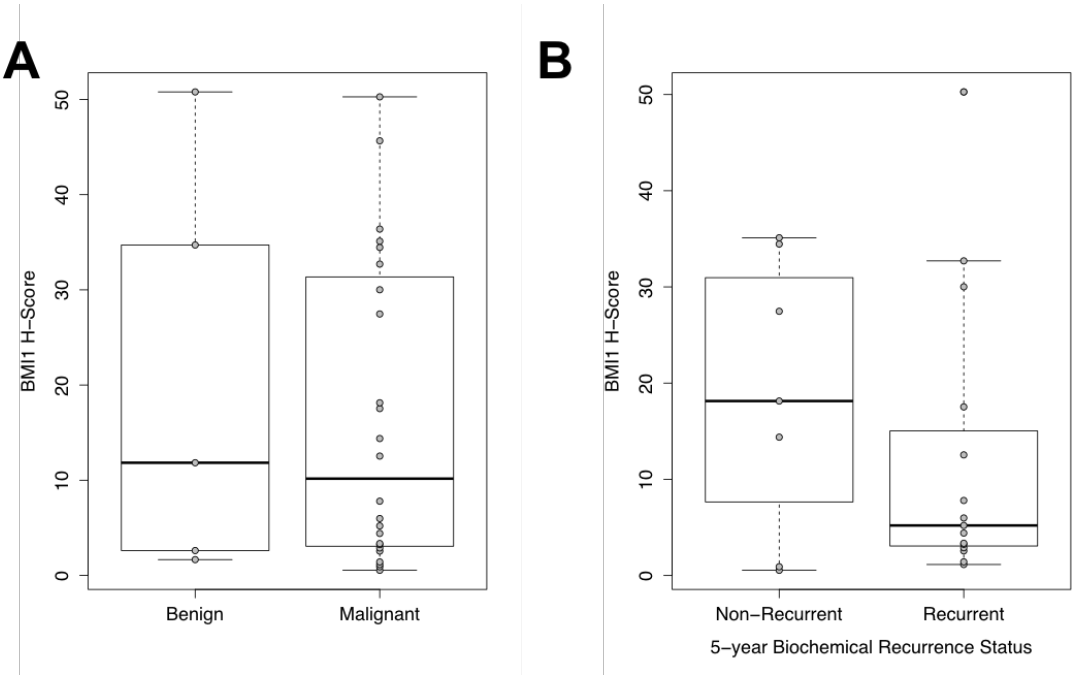


Figure 4.8: Quantification of BMI1 nuclear staining in benign prostate and prostate adenocarcinoma tissue.

IHC staining of BMI1 was quantified in the Bath cohort using the H-Score for nuclear stains. Boxplots of BMI1 nuclear H-score were compared between benign (n=5) and malignant prostate tissue (n=24) (A), showing no significant difference between the two groupings. 5-year recurrence-free (n=7) and recurrent patients (n=15) were also compared for BMI1 nuclear H-Score (B), where there was no statistically significant difference, although there was a trend towards decreased BMI1 expression in recurrent tumours. Data represent the mean of five random fields of view per sample. Two-sample t-tests were conducted to determine statistical significance for each set of conditions.

4.2.3.10 ZSCAN4 is overexpressed in a subset of prostate carcinomas

To continue to characterise the expression of stem cell markers in prostate carcinoma and their association with outcome, expression of the stem cell-related transcription factor ZSCAN4 was investigated by IHC. Since ZSCAN4 expression has not been reported in the prostate before, a positive control was required. ZSCAN4 expression is generally thought to be under tight spatiotemporal regulation during development, being present almost exclusively at the 2-cell stage of embryonic development (Falco *et al.*, 2007) and expressed in embryonic stem cells (Storm *et al.*, 2014). Therefore, positive control tissue for the IHC experiments would be difficult to obtain. Having previously identified that ZSCAN4 was transcribed at an mRNA level in placental tissue (Chapter 3.2.3.4), we hypothesised that human placental tissue would express ZSCAN4 protein and could serve as a positive control tissue. Normal human placenta was stained alongside benign and malignant prostate samples, and the Bath cohort was scored using the H-Score for nuclear antigens.

When stained with an anti-ZSCAN4 antibody, normal prostate tissue either expressed ZSCAN4 at low frequencies in the nuclei of glandular epithelium, either without diffuse glandular cytoplasmic staining (Figure 4.9A) or in its presence (Figure 4.9B). Many prostate carcinoma samples also share this pattern, displaying either rare ZSCAN4+ nuclei (Figure 4.9E) or none at all (Figure 4.9F). Glandular cytoplasmic staining of ZSCAN4 tended to be diffuse and weak to moderate intensity in tumour epithelium, and was not scored for further analysis. However, a subset of prostate carcinoma samples reacts more strongly positive for ZSCAN4 (Figure 4.9C and D, arrows), with 7 of 27 evaluable tumours (25.9%) having an H-score>=15 (Figure 4.10A). On average, malignant prostate tissues displayed significantly higher nuclear H-Scores compared to benign samples, with mean H-Scores of 11.1 and 4.1 respectively ($p=0.0342$) (Figure 4.10A). Although only a small proportion of the patients had an elevated burden of ZSCAN4-expressing tumour cells, this finding suggests that elevated ZSCAN4 expression is a tumour-specific phenomenon but is restricted to a subset of cases. There was no significant difference in ZSCAN4 expression between low Gleason grade ($\leq$3+4) and high Gleason grade ($\geq$4+3) tumours ($p=0.5354$).

Consistent with previous bioinformatics analyses at the mRNA level in Chapter 3, anti-ZSCAN4 antibody signal was detected in the nuclei of enlarged cells of the decidual

layer of the placenta (Figure 4.9G, arrows). A weak to moderate, diffuse, cytoplasmic staining reaction was also observed in this area. Positive staining in the expected compartment in the placenta provided extra confidence in the prostate immunohistochemistry. Importantly, no primary antibody negative controls in the placenta suggest that the detection method does not produce any background staining reactions in this tissue (Figure 4.9H).

4.2.3.11 ZSCAN4 is not significantly associated with 5-year biochemical recurrence status

The rate of disease recurrence following primary therapy is approximately 15-30% (Pound *et al.*, 1999; Stephenson *et al.*, 2006) and the frequency of evaluable tumours with ZSCAN4-overexpressing phenotypes (H-score>=15) falls within this range, at 25.9%. We used the clinical metadata to determine whether there was an association between ZSCAN4 staining and disease recurrence. As before, data were dichotomised based on a cut-off of 5 years from the date of diagnosis to biochemical recurrence. There was no significant difference between the scored levels of ZSCAN4 reactivity in recurrent patient tissues compared to non-recurrent ($p=0.636$) (Figure 4.10B). While 4 of 17 patients (23.5%) in the recurrent group had a high burden of ZSCAN4-positive tumour cells (H-Score >= 15), this was also the case for 3 of 10 non-recurrent patients (30%). Taken together, the data do not suggest that there is an association between ZSCAN4 reactivity and patient outcome within the first five years of treatment.

4.2.3.12 Assessment of ZSCAN4 nuclear staining is satisfactorily reproducible between two observers

To further validate the reproducibility of the H-Scoring system, we compared the scores observed and calculated by the author, BS, to those of a certified pathologist, JRM, for a subset consisting of 16 samples: 2 benign adjacent tissues and 14 prostate tumours. Inter-observer reproducibility was assessed by comparing paired scores for each image from a total of 77 images, and calculating the Pearson's product-moment correlation coefficient for the scores of both independent observers. Scoring between BS and JRM had a moderately strong correlation of 0.805 (95% CI: 0.709-0.872), which was statistically significant (two-tailed, $p<2.2e-16$). This correlation suggests that agreement between observers was good, as shown in Figure 4.10C, but scores became more variable as the H-Score increased, suggesting a higher inter-observer error rate in judging stronger intensities and/or higher proportions of stained cells (Figure 4.10D).

Overall, the good correlation suggests that the scoring procedure for assessing nuclear ZSCAN4 levels was suitably reproducible between observers.

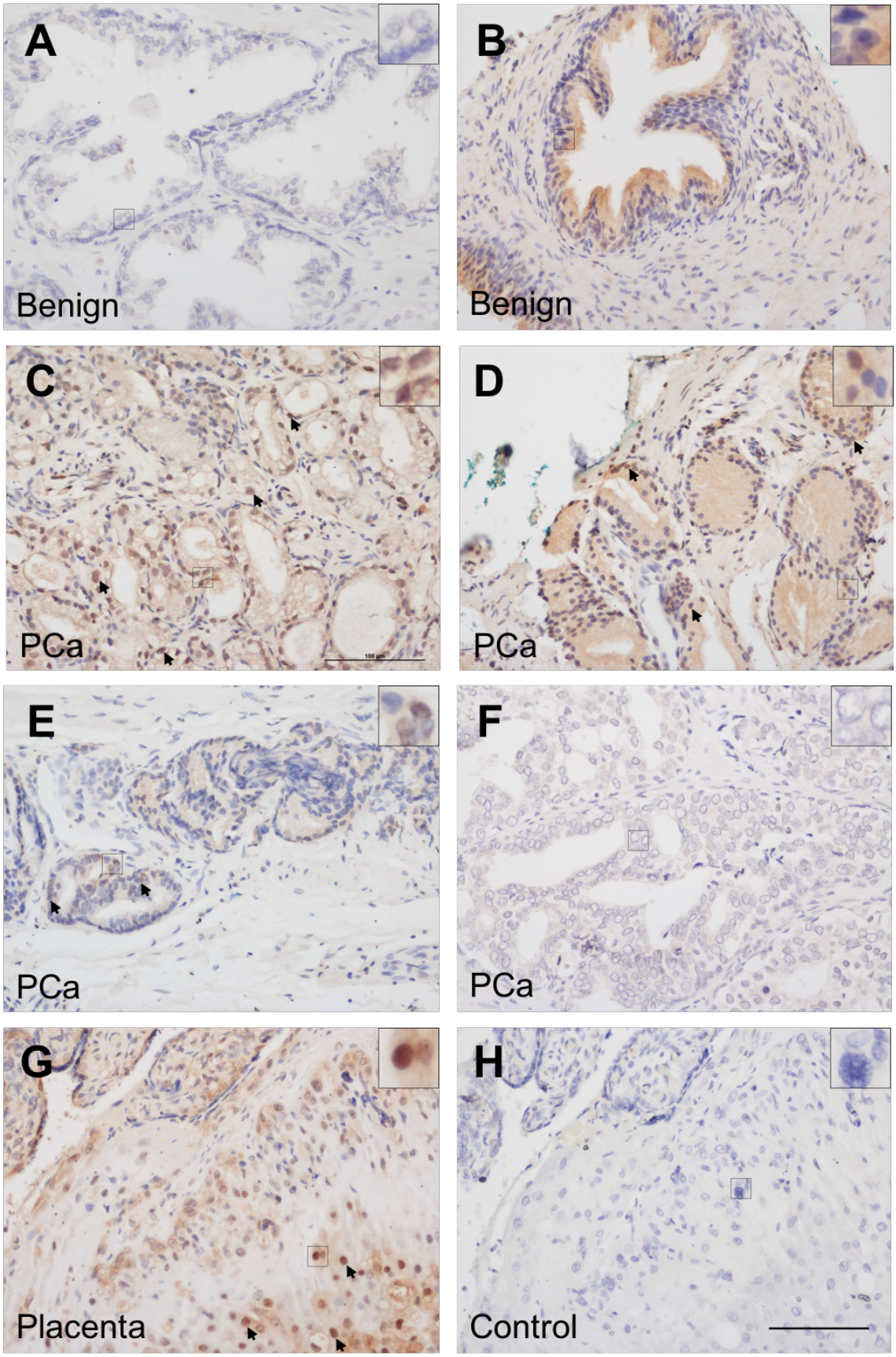


Figure 4.9: Anti-ZSCAN4 IHC in benign and malignant prostate tissue.

ZSCAN4 nuclear staining is low or negative in benign prostate tissue (A), with uniform diffuse staining of epithelial cytoplasm occasionally observed (B). Conversely, proportions of ZSCAN4+ nuclei varied widely in malignant samples, with a subset of prostate carcinoma samples (PCa) displaying a high burden of ZSCAN4+ epithelial nuclei (C, arrows). Malignant samples with intermediate (D, arrows), low (E, arrows) or negative nuclear staining (F) were observed. Placenta serves as a positive control for ZSCAN4 staining, showing ZSCAN4+ nuclei in the decidual layer (G, arrows). Negative controls are free of background staining, as shown by the no primary antibody control in placental tissue (H). Antibody binding is visualised with DAB (brown) and nuclei are counterstained with haematoxylin (blue). Representative fields of view are shown at 20x magnification, with representative epithelial cells shown inset and representative positive cells indicated with arrows. Scale bar – 100μm.

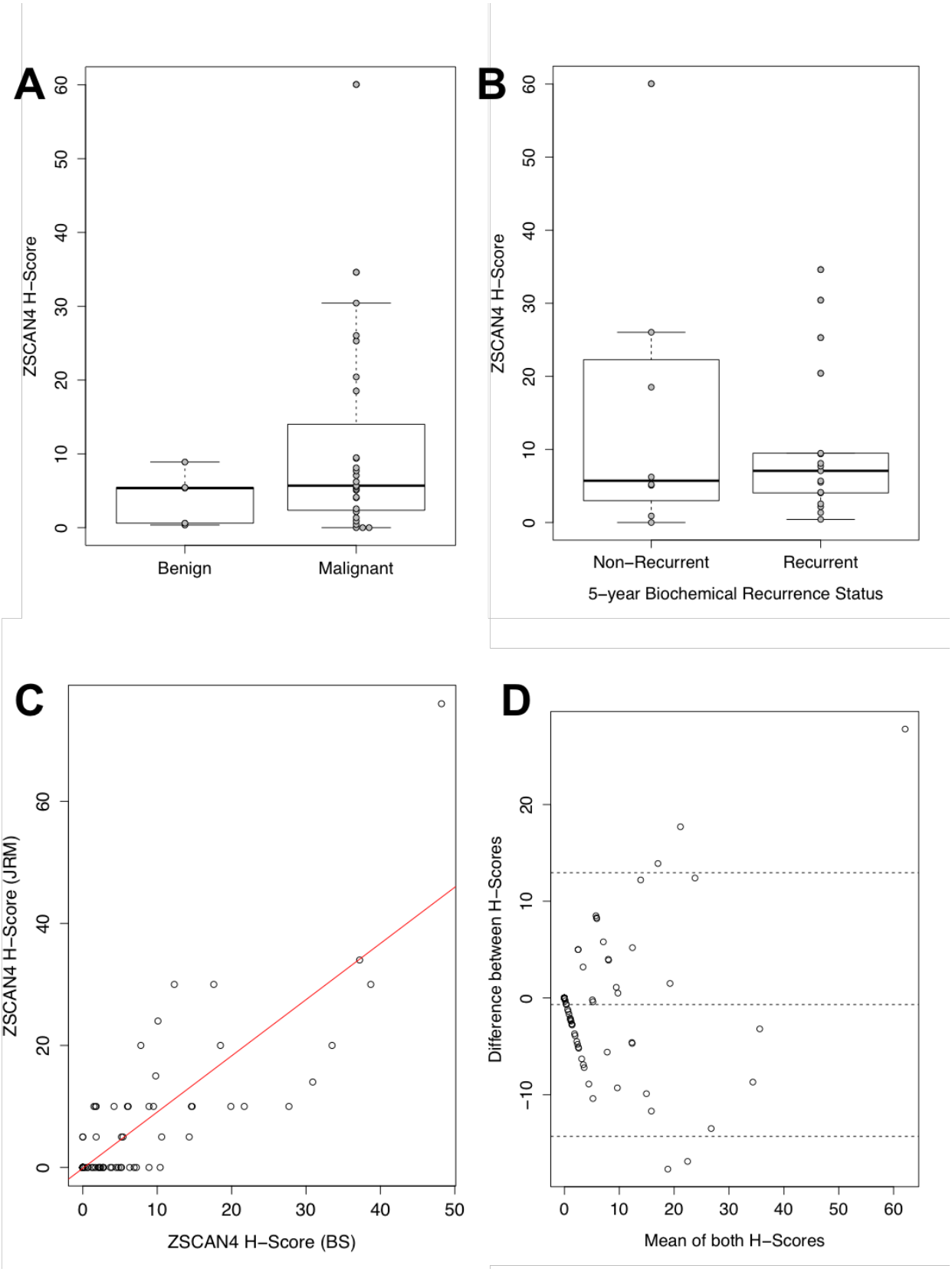

A
ZSCAN4 H−Score
0
10
20
30
40
50
60
Benign
Malignant
B
ZSCAN4 H−Score
0
10
20
30
40
50
60
Non−Recurrent
Recurrent
5−year Biochemical Recurrence Status
C
ZSCAN4 H−Score (JRM)
0
20
40
60
0
10
20
30
40
50
ZSCAN4 H−Score (BS)
D
Difference between H−Scores
−10
0
10
20
0
10
20
30
40
50
60
Mean of both H−Scores

Figure 4.10: Quantification and Reproducibility assessment of ZSCAN4 nuclear staining.

IHC staining of ZSCAN4 was quantified in the Bath cohort using the H-Score for nuclear stains. Boxplots of ZSCAN4 H-Scores were compared between benign (n=5) and malignant prostate tissue (n=27) (A), showing significantly higher ZSCAN4 H-Scores in malignant tissue (p=0.0342). 5-year recurrence-free (n=8) and recurrent (n=17) patients were also compared for ZSCAN4 nuclear H-Scores (B), where there was no statistically significant difference. Data represent the mean of five random fields of view per sample. Two-sample t-tests were conducted to determine statistical significance for each set of conditions in A and B. A subset of 16 samples were scored by both the author (BS) and a certified pathologist (JRM) and the resulting paired scores were compared in (C); There was a statistically significant, moderately strong Pearson's correlation between the paired scores of both observers ($R=0.805$, $p<2.2e-16$). The differences and means of paired H-Score values are visualised in a Bland-Altman plot (D), showing that most paired values lie ± two standard deviations from the mean (dashed lines).

4.2.3.13 RS1 is expressed in both benign and malignant prostate tissue

We previously established RS1 as being a candidate differentially expressed gene between relapsing and non-relapsing prostate carcinoma using differential expression analysis of archival prostate tumour cDNA microarray data (Chapter 3.2.2). To test this hypothesis in a cohort of patients, we performed immunohistochemistry with an anti-RS1 antibody. As RS1 is a secreted protein expressed in the retina and has not been studied in any other tissue type, a positive control tissue was not feasible and its expression was studied in benign and malignant prostate only. In benign tissue, anti-RS1 antibody signal localised to the cytoplasm of glandular epithelial cells, with heterogeneous staining of epithelial nuclei, and considerable heterogeneity between cases. Nuclear staining was unexpected given the role of RS1 as a secreted protein. Within normal cases, 1/5 tumours strongly stained for nuclear RS1 and 2/5 showed moderate staining for cytoplasmic RS1 (Figure 4.11A). The remaining tumours had weak to negative RS1 staining in both the nucleus and cytoplasm (Figure 4.11B). Within the malignant prostate, nuclear staining ranged from intense and heterogeneous (Figure 4.11C and D), to weak and infrequent (Figure 4.11E) and negative (Figure 4.11F). Cytoplasmic staining equally ranged from intermediate (Figure 4.11C) to weak (Figure 4.11D) and negative (Figure 4.9E and F). Negative controls confirmed that the detection protocol was background-free (Figure 4.11G).

RS1 staining was scored using the nuclear H-score and cytoplasmic score for intensity only. Interestingly, paired benign and malignant tissue from the same patient had high levels of RS1 nuclear reactivity (Figure 4.11 A and C respectively), with a higher H-Score of 126.6 in the benign tissue and 51.5 in malignant tissue respectively. However, overall there was no significant difference in RS1 nuclear (Figure 4.12A) or cytoplasmic reactivity scores (Figure 4.12B) between benign and malignant prostate tissue ($p=0.5327$ and 0.7006 respectively). There was also no significant difference between low Gleason grade ($\leq 3+4$) and high Gleason grade ($\geq 4+3$) tumours with respect to RS1 nuclear or cytoplasmic expression ($p=0.8203$ and 1 respectively)

4.2.3.14 RS1 has higher mean expression in 5-year biochemical recurrent samples

RS1 staining was assessed at both the nuclear and cytoplasmic levels, and compared between recurrent and recurrence-free patients after 5 years of retrospective follow-up as before. There was no significant difference between nuclear H-Scores in recurrent and non-recurrent patient groupings (Figure 4.12C) ($p=0.4954$). There appeared to be a trend towards increased RS1 cytoplasmic scores in recurrent prostate tissue (Figure

4.12D), although this was not significant (p=0.4506). This appears consistent with findings in Chapter 3, which identified increased RS1 mRNA expression in patients with subsequent recurrence. The trends of increasing RS1 cytoplasmic reactivity would benefit from further investigation in a larger cohort of patients, especially in light of other findings from Chapter 3 supporting the hypothesis, such as: increased mRNA expression in prostate tumours of recurrent patients on microarray analysis; and the finding of RS1 amplification in 20-30% of patients with metastatic and castration-resistant prostate carcinoma.

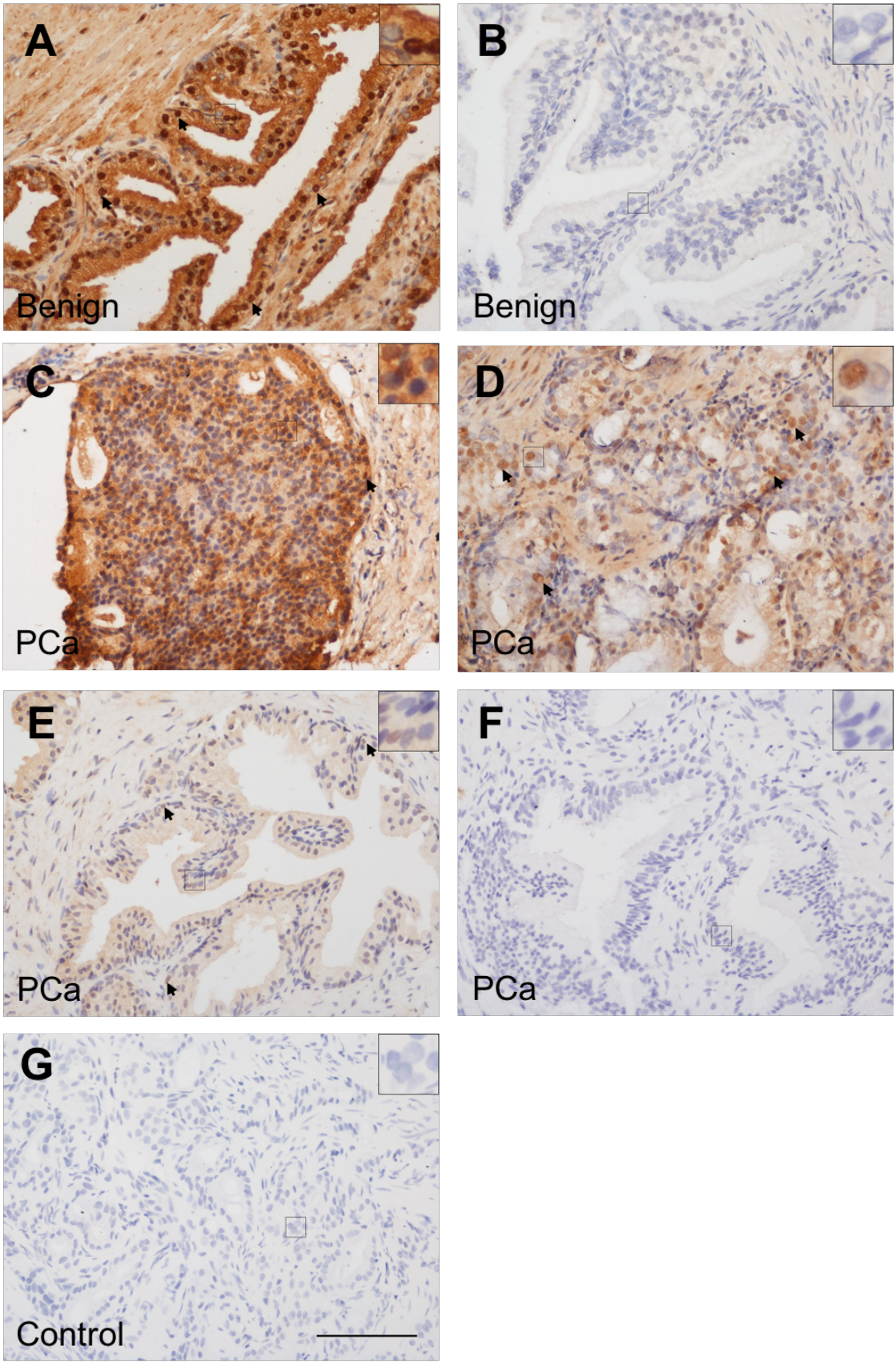


Figure 4.11: Anti-RS1 IHC in benign and malignant prostate tissue.

RS1 localised to the nucleus and cytoplasm of benign prostate tissue, with variation in both staining patterns ranging from frequent but heterogeneous positive nuclei and intermediate cytoplasmic staining (A) to no nuclear staining and no cytoplasmic staining (B). Malignant prostate samples (PCa) display a similar expression pattern to benign samples, with large variation in both the cytoplasmic and nuclear stain of epithelial cells: intermediate intensity cytoplasmic staining and intermediate frequency nuclear staining (C); paired tumour sample from sample in A, with weak cytoplasmic staining and similarly high frequency nuclear staining (D); weak cytoplasmic staining and low frequency nuclear staining (E); negative staining for both compartments (F). Negative controls are free of background staining, as shown by the no primary antibody control in prostate carcinoma tissue (G). Antibody binding is visualised with DAB (brown) and nuclei are counterstained with haematoxylin (blue). Representative fields of view are shown at 20x magnification, with representative epithelial cells shown inset and representative positive cells indicated with arrows. Scale bar – 100μm.

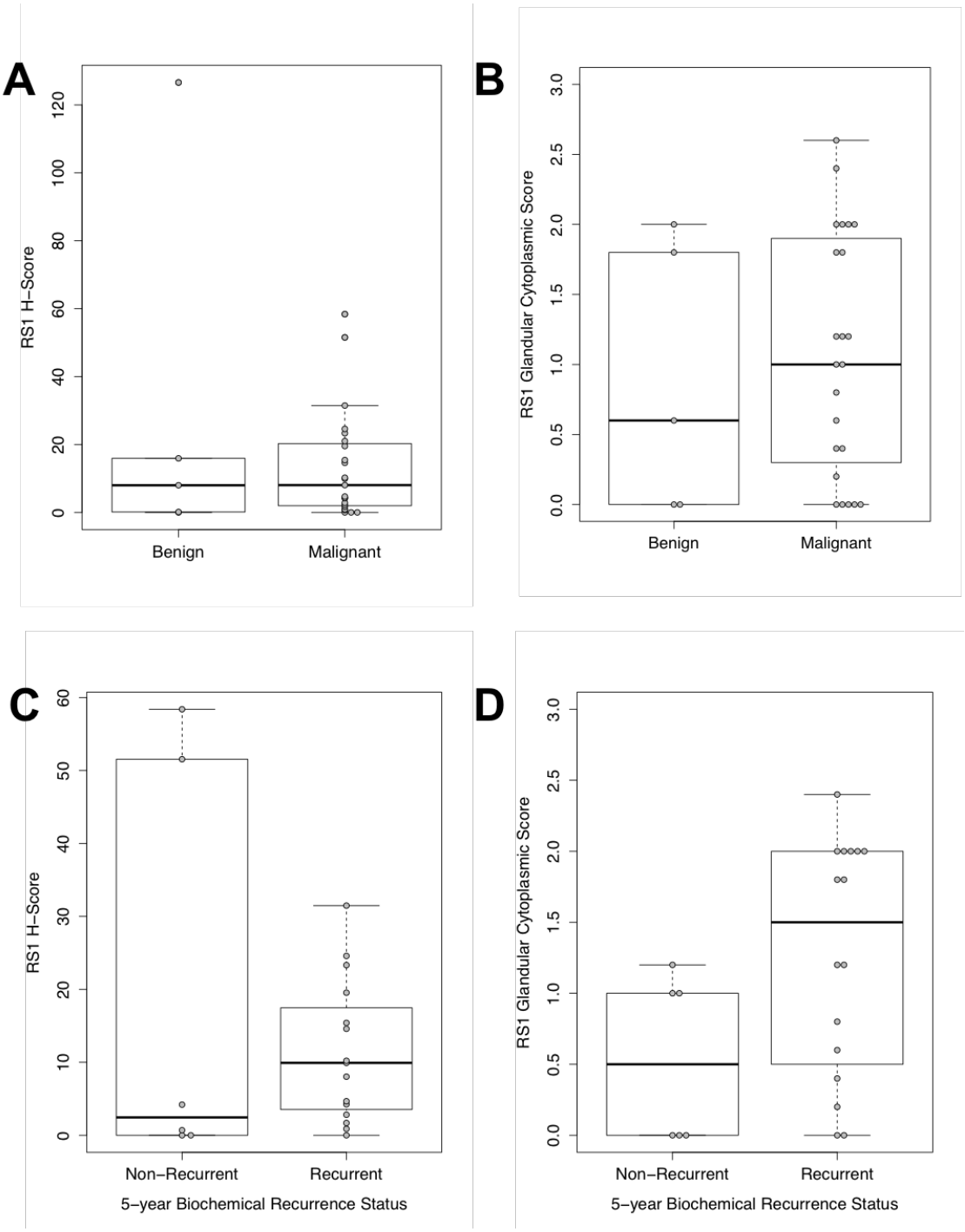

A
RS1 H–Score
0
20
40
60
80
100
120
Benign
Malignant
B
RS1 Glandular Cytoplasmic Score
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Benign
Malignant
C
RS1 H–Score
0
10
20
30
40
50
60
Non–Recurrent
Recurrent
5–year Biochemical Recurrence Status
D
RS1 Glandular Cytoplasmic Score
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Non–Recurrent
Recurrent
5–year Biochemical Recurrence Status

Figure 4.12: Quantification of RS1 nuclear and cytoplasmic staining in benign and malignant prostate tissue.

Staining was quantified using the H-Score and Cytoplasmic score. Boxplots of RS1 (A) nuclear H-Score and (B) Cytoplasmic score compared between benign (n=5) and malignant prostate tissue (n=23), showing no difference in RS1 levels between the two groupings. 5-year recurrence-free (n=6) and recurrent patients (n=15) were also compared for RS1 (C) nuclear H-Score and (D) Cytoplasmic score, where there was no significant difference, although a trend towards increased RS1 cytoplasmic staining was observed in recurrent patients. Data represent the mean of five random fields of view per sample. Two-sample t-tests were conducted to determine statistical significance for each pair of conditions.

4.2.4 Analysis of AR Expression in Prostate Tumours

Expression of the AR has been shown to be more heterogeneous in prostate tumours than in benign prostate epithelium (Barboro *et al.*, 2014), and studies suggest that AR expression (Donovan *et al.*, 2008) and its downstream targets (Hendriksen *et al.*, 2006) could be an important prognostic factor for patients with clinically localised prostate cancer. AR expression is positively correlated with other poor prognostic factors, such as tumour stage and Ki-67 labelling index for proliferating cells (Goltz *et al.*, 2015). Therefore, a potential biomarker which displays a correlation with androgen receptor status may itself be subject to regulation by androgen signalling, and could also vary due to subsequent anti-androgen therapies. To assess whether selected candidate biomarkers co-varied with AR expression in the Bath cohort, immunohistochemistry was performed using a monoclonal anti-AR antibody. AR expression was analysed independently with respect to clinical data, before subsequent analysis with the candidate biomarkers.

4.2.4.1 AR expression is heterogeneous, especially in prostate carcinoma

AR expression data was obtainable 29 of 35 specimens (82.9%), including 5 adjacent benign prostate biopsies. Insufficient material was the only grounds for specimen exclusion. AR signal was predominantly localised to epithelial nuclei in both benign (Figure 4.13A and B, arrows) and malignant prostate tissue (Figure 4.13C-E). AR expression was heterogeneous even in benign tissue, with a salt-and-pepper pattern of nuclear AR staining in the luminal prostate epithelium (Figure 4.13A). Interestingly, 2 of 5 benign samples had an H-Score of less than 10, indicating a very low proportion of cells expressing detectible AR protein (Figure 4.13B). Malignant prostate samples also tended to display this heterogeneous expression of AR, but appeared to have greater variation in the frequency and intensity levels, as represented by a greater range of H-Score values (Figure 4.14A). Malignant expression patterns varied considerably, from widespread moderate intensity nuclear staining (Figure 4.13C) to heterogeneous moderate nuclear staining (Figure 4.13D); weak and heterogeneous nuclear staining (Figure 4.13E); and an absence of nuclear AR staining (Figure 4.13F). No primary antibody controls indicated that the detection system is background-free (Figure 4.13G). On average, benign tissues displayed a trend towards greater AR H-Scores than malignant tissues (29.0 and 23.4 respectively) (Figure 4.14A), but this was not statistically significant ($p=0.6887$) and would require more benign samples to test

conclusively. Additionally, there was no significant difference in AR expression between low Gleason grade (≤3+4) and high Gleason grade (≥4+3) tumours (p=0.1936).

4.2.4.2 AR expression is not associated with 5-year biochemical relapse status

Following comparisons between benign and malignant tissue, AR expression was compared between recurrent and non-recurrent patient groups. 22 patients with sufficient follow-up information were assessed for AR status, including 16 recurrent and 6 non-recurrent patients after 5 years of retrospective follow-up (Figure 4.14B). There was a trend towards reduced mean H-Score of AR in recurrent patient tissues (29.0 recurrent and 23.4 non-recurrent), but this was not statistically significant (p=0.6887).

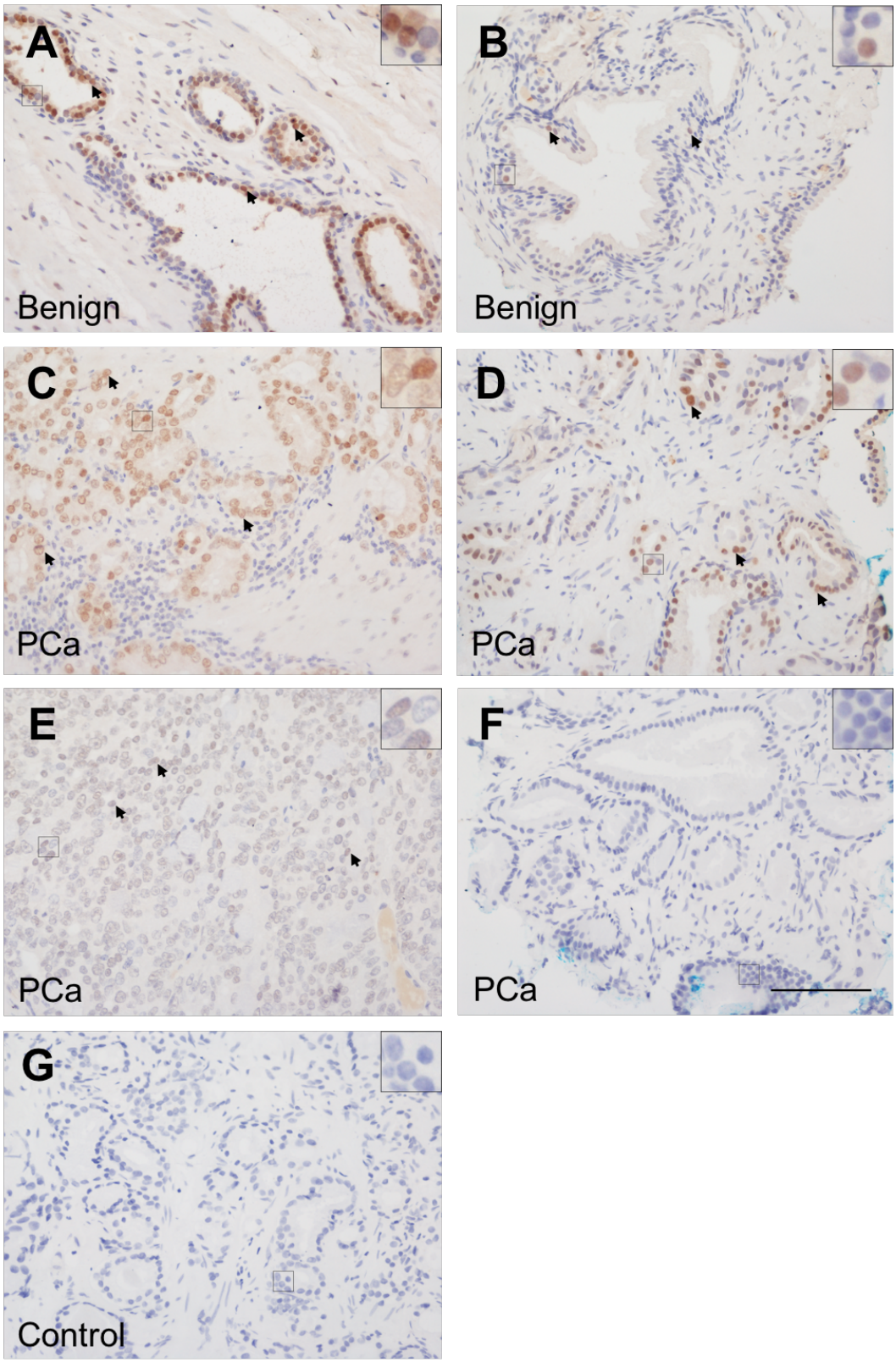

A
Benign
B
Benign
C
PCa
D
PCa
E
PCa
F
PCa
G
Control

Figure 4.13: Anti-AR IHC in benign and malignant prostate tissue.

AR is expressed heterogeneously in the luminal epithelial nuclei in benign tissue (A, arrows), although 2/5 benign prostate samples have a very low frequency of AR staining (B). Prostate tumours (PCa) also show heterogeneous but predominantly nuclear staining, ranging in frequency from widespread (C) to mixed (D) to weak, low frequency staining (E) and complete absence of AR staining (F). Negative controls are free of background staining, as shown by the no primary antibody control in prostate carcinoma tissue (G). Antibody binding is visualised with DAB (brown) and nuclei are counterstained with haematoxylin (blue). Representative fields of view are shown at 20x magnification, with representative epithelial cells shown inset and representative positive nuclei indicated with arrows.

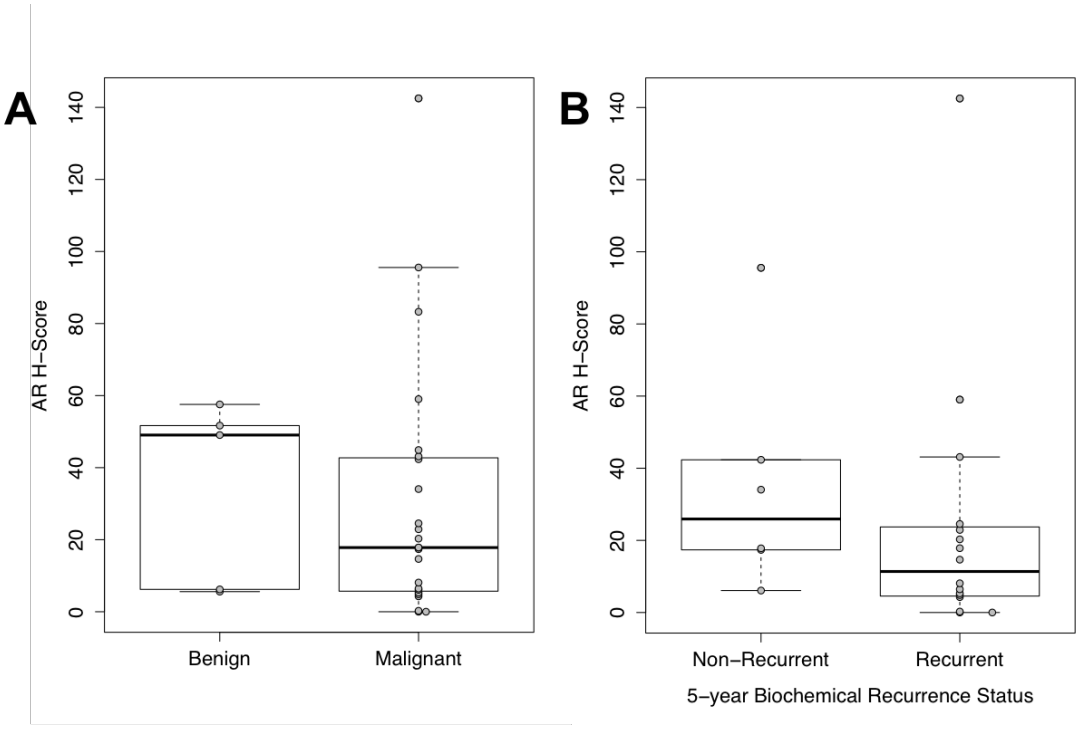


Figure 4.14: Quantification of nuclear AR staining in benign and malignant prostate tissue by IHC.

AR staining was quantified using the H-Score for nuclear staining. Boxplots comparing AR nuclear H-Scores in (A) benign (n=5) and malignant (n=24) tissue samples; and (B) 5-year recurrence-free (n=6) and recurrent (n=16) samples. There was no significant difference between AR H-Scores in benign and malignant tissue, although a trend towards decreased AR cytoplasmic staining was observed in recurrent patients. Data represent the mean of five random fields of view per sample. Two-sample t-tests were conducted to determine statistical significance for each pair of conditions.

4.2.4.3 Assessment of correlation between AR and candidate biomarker expression

Having established the expression patterns of AR in malignant prostate tissue, the nuclear H-Scores were tested for correlation with the scores from the candidate protein biomarkers and pre-existing clinical biomarkers such as age and PSA levels at diagnosis (Table 4.3). Due to the lack of patients in the N1 and M1 (N=2/27 each), and the wide dispersal of a small number of samples over a wide range of T categories, we were unable to assess any relationships between biomarker expression and TNM staging in this small cohort.

Among the Bath cohort, nuclear AR expression was not associated with age or PSA levels at diagnosis. The only candidate biomarker that was significantly correlated with AR expression was RS1, and only the glandular cytoplasmic stain was significant with a moderately negative correlation of -0.454 (p=0.04459) (Figure 4.15A). Interestingly, nuclear expression of RS1 had no significant correlation with AR expression (Figure 4.15B). A summary of the correlations between AR and potential biomarkers is provided in Table 4.4.

Table 4.3: Correlations between candidate biomarkers and clinical variables and AR H-Scores.

Biomarker	*Correlation Coefficient*	*95% CI Lower*	*95% CI Upper*	*p-Value*
RS1 Glandular Cytoplasmic Score	-0.454	-0.746	-0.014	*0.045*
BMI1 H-Score	0.339	-0.108	0.672	*0.133*
Age at Diagnosis	0.326	-0.111	0.657	*0.139*
ALDH7A1 H-Score	0.346	-0.144	0.700	*0.159*
PCSP Total Area	0.256	-0.198	0.619	*0.263*
PCSP Cell Number	-0.211	-0.589	0.243	*0.359*
PSA at Diagnosis	-0.190	-0.603	0.304	*0.451*
RS1 H-Score	0.162	-0.315	0.574	*0.507*
SDC1 Glandular Cytoplasmic Score	0.149	-0.302	0.546	*0.518*
ALDH7A1 Glandular Cytoplasmic Score	0.154	-0.338	0.579	*0.543*
ZSCAN4 H-Score	0.109	-0.351	0.526	*0.648*
Nestin Glandular Cytoplasmic Score	-0.077	-0.503	0.378	*0.747*

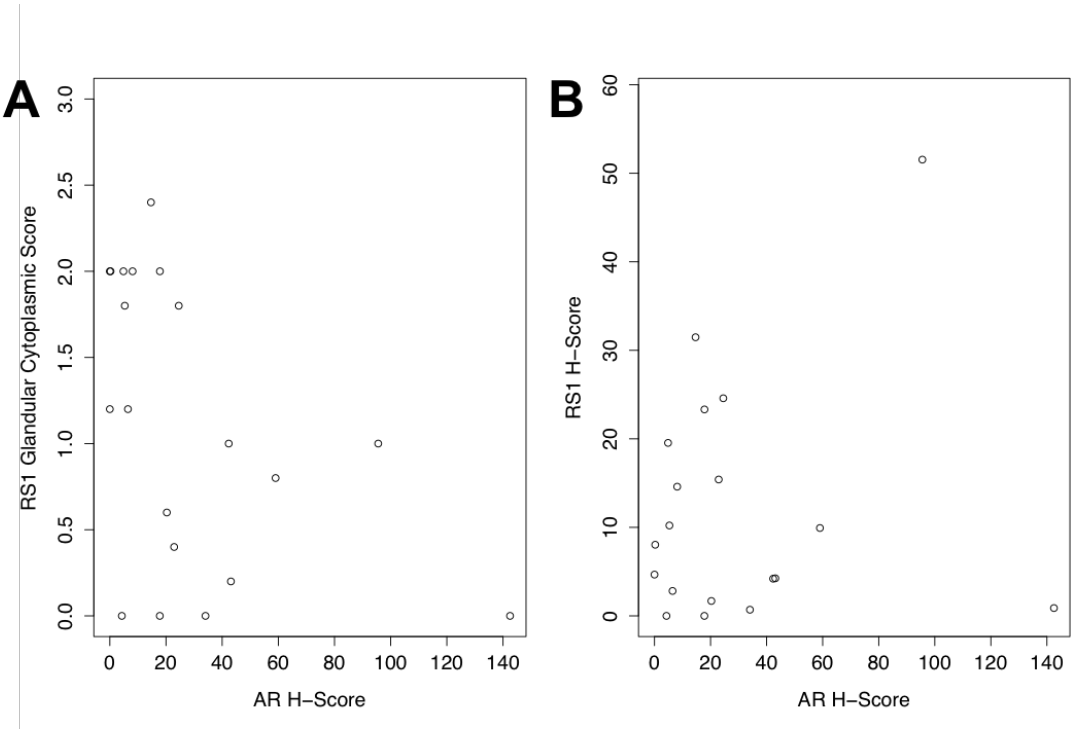


Figure 4.15: Plots of RS1 cytoplasmic and nuclear scores against paired AR nuclear H-Scores.

AR nuclear H-Scores were plotted against RS1 Cytoplasmic Scores and nuclear H-Scores. RS1 glandular Cytoplasmic Scores displayed a significant and moderately negative correlation with AR H-Scores (R=-0.454, p=0.04459) (A), whereas RS-1 nuclear H-Scores showed no statistically significant correlation with AR expression (B). n=26 samples (21 malignant, 5 benign).

Table 4.4: Summary of findings for each biomarker in this chapter.

	Comparisons:		
Biomarker	Benign vs. Malignant	Non-Recurrent vs. Recurrent	AR Expression Correlation
ALDH7A1:			
Nuclear	No trend	No trend	None
Cytoplasmic	No trend	No trend	None
SLC31A1:	Excluded from analysis due to strong ubiquitous staining		
MUC1-C:	Excluded from analysis due to no detectable staining		
Nestin:			
Cytoplasmic	No trend	No trend	None
BMI1:			
Nuclear	No trend	**Lower in recurrent** (not significant)	None
ZSCAN4:			
Nuclear	Higher in subset of malignant (p=0.0342)	No trend	None
RS1:			
Nuclear	No trend	No trend	None
Cytoplasmic	No trend	Higher in recurrent (not significant)	Modest Negative (R=-0.454)

4.3 Discussion

In the current chapter, the expression of candidate biomarkers in benign and malignant prostate tissue has been addressed, and recurrent and non-recurrent prostate tumours have been compared for expression of these markers. A summary of the experimental findings of this chapter is presented in Table 4.4. BMI1 was found to be expressed in prostate tumours but was found to be reduced in recurrent tumours, although this was not significant and requires a larger sample size to verify. The stem cell marker ZSCAN4 was found to be expressed in a subset of prostate tumours. Additionally, RS1 expression in the cytoplasm appears to be correlated with androgen receptor status in the primary tissue of patients with prostate carcinoma.

4.3.1 ALDH7A1 is not associated with biochemical recurrence but is found in the nucleus of benign and malignant prostate tissue

Studies on ALDH7A1 have focused on its role in invasiveness and metastasis, with functional assays being conducted in vitro and in xenograft models in vivo (van den Hoogen *et al.*, 2011). It is expressed in both human prostate tumours and matched bone metastases (van den Hoogen *et al.*, 2010). However, the results of this chapter suggest that ALDH7A1 expression is unable to distinguish between patients that will later relapse and those that will remain in remission. It is possible that it plays a role in later stages of tumour progression, such as invasion and metastasis, and therefore is not prognostically relevant in primary tissue at the time of therapy.

Interestingly, van den Hoogen *et al*. (2010) only reported staining at the cytoplasmic level in their IHC studies. I report both cytoplasmic and nuclear staining in benign and malignant prostate tissue. Nuclear localisation of ALDH7A1 has been observed before in mouse and human tissues by immunohistochemistry, and both nuclear export and nuclear localisation signals have been found in the human protein (Brocker *et al.*, 2010). The function of nuclear ALDH7A1 is currently unknown and could be a subject for future investigation.

4.3.2 SLC31A1 shows strong ubiquitous IHC staining and was excluded

SLC31A1 was identified as overexpressed in recurrent patients in a microarray analysis in chapter 3. During optimisation, SLC31A1 was excluded from the panel of potential biomarkers due to the presence of a strong background staining in the stroma and a ubiquitous and strong staining in the prostate epithelium, both of which would hinder

interpretation of the biomarker by IHC. This strong and ubiquitous staining pattern in both prostate epithelium and stroma could reflect a requirement for copper homeostasis in all cells of the prostate gland (Jong and McKeage, 2014). Conversely, one study conducted by Holzer *et al.* suggests that the protein is not expressed at the protein level in benign prostate tissue or in prostate carcinoma (Holzer *et al.*, 2006), although this is unconfirmed due to reports of SLC31A1 expression in VCaP and LNCaP prostate cancer cell lines, and the dependence of benign prostate and prostate cancer cell lines on copper (Safi *et al.*, 2014). Expression of SLC31A1 was identified at the mRNA level during this thesis, but this does not necessarily translate into upregulated protein levels and it is possible that the antibody staining results are a non-specific background staining effect. However, this biomarker may still be of use with further optimisation of IHC conditions and testing in more samples, and another antibody raised to this protein has the potential to provide more satisfactory staining with less background. Further investigation of SLC31A1 expression in prostate cancer would ideally require IHC staining with another antibody to confirm the reliability of these results.

4.3.3 MUC1-C is not detected by IHC in prostate cancer samples

Initial optimisation of the MUC1-C antibody for staining indicated that it produced no IHC signal on any of the four tissue samples studied. While this does not exclude the potential for its expression in a subset of carcinoma samples, this led to its exclusion from the panel of potential biomarkers, particularly due to the lack of defined positive control material to confirm staining specificity. There is also the possibility of MUC1-C expression being acquired only after therapy, rather than being expressed in a pre-existing prostate cancer cell population. Androgen-independent prostate cancer cell lines that lack the androgen receptor, such as PC-3, appear to express MUC1-C, whereas androgen-sensitive cell lines which express AR do not (Joshi *et al.*, 2009). MUC1-C expression in AR-expressing cell line LNCaP causes AR expression to be reduced and confers the ability to grow in the presence of the AR antagonist bicalutamide (Rajabi *et al.*, 2012). However, prostate cancer cell lines are nearly all derived from metastatic lesions, and the expression of MUC1-C has yet to be confirmed in human prostate tumours. Staining a tissue microarray containing both primary tumour samples and metastatic and castration-resistant prostate carcinoma samples would help to confirm the relevance of previous *in vitro* studies to a patient setting.

4.3.4 Nestin is not associated with biochemical recurrence

Nestin was expressed weakly in the glandular compartment of benign prostate epithelia, and expression was detected in the same compartment in malignant prostate epithelia with intensities ranging from intermediate to negative. Kleeberger *et al.* found that Nestin-positive endothelial cells were scattered in hormone-naïve prostate epithelia, but reported no Nestin staining in tumour cells of hormone-naïve tumours, whereas Nestin-positive tumour cells were widespread in castration-resistant prostate tumours (2007b). We did not observe a staining pattern that would be consistent with endothelial cells as reported by Kleeberger, although we identified scattered stromal cells which infrequently stained weakly with anti-Nestin antibody and are of unconfirmed identity. Additionally, we found no association between staining levels and 5-year recurrence status. It might be that Nestin expression is induced as a result of hormone therapy, a finding that is supported by several groups (Kleeberger *et al.*, 2007b; Pfeiffer *et al.*, 2011). Indeed, another study found that Nestin expression had no association with clinical variables on its own, but rather only when combined with proliferative marker Ki-67 (Gravdal *et al.*, 2009). As the proliferation status of this cohort has not been determined, the findings of Gravdal *et al.* cannot be verified. Additionally, the Bath cohort did not include paired post-treatment biopsies, which would be required for verifying the hypothesis of post-treatment Nestin expression. Overall, the data in this chapter cannot support a role for Nestin in the prognostication of prostate carcinoma at point of primary therapy.

4.3.5 BMI1 may be negatively associated with biochemical recurrence

Others have reported the utility of BMI1 as a serum prognostic biomarker which positively correlates with tumour stage and PSA levels (Siddique *et al.*, 2013c). This same group also demonstrated a stromal staining of BMI1 by IHC, which I do not observe in this study. Antibodies have individual specificities and sensitivities to their target antigens which also depend on the immunohistochemistry protocol employed, making the comparison of studies performed with different antibodies to the same protein difficult. Although the antibody used by Siddique et al. was from the same manufacturer, it was not possible to confirm whether the same antibody clone was used in the current work. However, consistent with a previous study (van Leenders *et al.*, 2007), BMI1 localised to epithelial nuclei in both benign and malignant prostate tissue, which helps to verify our observed staining patterns.

Van Leenders (2007) found that strong epithelial BMI1 staining was positively associated with biochemical recurrence. We have found that levels of nuclear BMI1 staining have no significant association with biochemical recurrence (Figure 4.6B). Conversely, the findings show a trend towards decreased BMI1 nuclear staining in recurrent patients compared to non-recurrent controls. There is evidence that BMI1 influences the progression and treatment response of prostate cancer. Reducing BMI1 levels in prostate cancer cell lines sensitises them to Docetaxel treatment (Crea *et al.*, 2011a; Siddique *et al.*, 2013a), possibly due to reduced antioxidant production (Crea *et al.*, 2011a) and the suppression of TCF4-driven expression of the anti-apoptotic protein BCL2 (Siddique *et al.*, 2013a). BMI1 is known to act as part of the PRC1 polycomb repressor complex to repress cyclin-dependent kinase inhibitors p21 and $p16^{Ink4a}$ in prostate cancer cell lines (Cao *et al.*, 2011), which would promote cell cycle progression. The PRC2 polycomb repressor complex epigenetically silences microRNAs which are responsible for downregulation of PRC1 complex members, including BMI1, in the DU145 prostate cancer cell line (Cao *et al.*, 2011). Expression of these microRNAs reduces cell proliferation, cell invasion and tumoursphere formation in DU145 cells *in vitro*, as well as reducing the expression of stem cell markers KLF4, SOX2 and c-Myc, suggesting that PRC1 activity is important for these phenotypes (Cao *et al.*, 2011). Indeed, suppressing the activity of the PRC2 complex by inhibition of its histone methyltransferase EZH2 reduces tumorigenicity and invasiveness in vitro and in mouse xenograft models (Crea *et al.*, 2011b). These lines of evidence suggest that the protein would contribute towards treatment resistance and tumour progression. Therefore, our results were surprising and a larger cohort of patients would be required to validate the observation of decreased BMI1 staining in the primary tissue of recurrent patients.

4.3.6 ZSCAN4 is expressed in a subset of prostate carcinomas, but is not associated with biochemical recurrence

ZSCAN4 was found to be expressed at very low levels in benign epithelium, with a subset of prostate carcinomas (25.9%) showing elevated expression which was predominantly nuclear (Figure 4.7 and 4.8A). Currently, there is very little known about the molecular function of ZSCAN4 and still less is known about it in tumour cells, with only two studies from Gollahon and Lee (2014, 2015) suggesting that it interacts indirectly with TRF1 to promote telomerase-independent elongation of telomeres in several cancer cell lines. This is the first study to demonstrate the expression of

ZSCAN4 in prostate cancer, and the fact that it is only present in a subset of prostate carcinomas suggests that it is fulfilling an unknown function in a subpopulation of patient tumours.

Given its previously demonstrated roles in promoting telomere elongation and self-renewal potential of embryonic stem cells (Storm *et al.*, 2014), it is possible that it is fulfilling the same functions in tumour cells. Telomere elongation would be more highly required in proliferating cell populations, and it would be worth investigating whether the proliferative index of tumours correlates with ZSCAN4 status. Although we could not demonstrate a relationship between ZSCAN4 expression and recurrence (Figure 4.8B), it is possible that ZSCAN4 is performing other interesting biological functions within tumour tissue that have yet to be investigated. For example, in mouse embryonic stem cells, ZSCAN4 expression is positively regulated by the p110α isoform of PI3K (Storm *et al.*, 2014), and PI3K-Akt signalling is known to be deregulated by deletions of PTEN in prostate cancer. The PTEN deletion status of the Bath Cohort has not been identified, and is another possible avenue for further characterisation of ZSCAN4's role in prostate cancer. However, this protein would need to be functionally characterised in primary prostate tumour cells or cancer cell lines to build on the data acquired here.

4.3.7 Cytoplasmic RS1 is expressed in benign and malignant prostate tissue and may be positively associated with biochemical recurrence

The secreted discoidin domain containing protein RS1 was found to be expressed in both benign prostate tissue and prostate carcinoma, with localisation to both the glandular cytoplasm and heterogeneously within nuclei (Figure 4.9 and 4.10A and B). This is unexpected given its extracellular localisation within the retina of mice (Ou *et al.*, 2015), where its functional role is believed to lie in cell adhesion, possibly mediated through its discoidin domain (Wu *et al.*, 2005). Interrogation of the protein sequence in the Uniprot database (Uniprot identifier: O15537) suggests that the RS1 protein has a signal sequence, meaning that it is destined for secretion. Therefore, it is unlikely that the nuclear staining reaction of the anti-RS1 antibody is a specific one.

Mutations in the protein cause splitting of the retinal layers, a disease known as X-linked retinoschisis after which the protein is named (Wu *et al.*, 2005). Given the lack of normal human retina as a positive control tissue in this study, it is difficult to confirm the specific binding of this antibody to prostate tissue. If the staining pattern is specific,

this is the first time that RS1 expression has been demonstrated outside of the eye, in benign prostate tissue and in prostate tumour tissue.

I observed a trend towards increased cytoplasmic staining intensity in patients that would go on to develop recurrent disease. I also observed a modest negative correlation between RS1 cytoplasmic staining and AR nuclear staining (Figure 4.13). It is possible that RS1 is upregulated as a result of reduced AR expression as the correlation suggests, or that AR represses RS1, as has been demonstrated for a number of genes in prostate cancer cell line C4-2B (Prescott *et al.*, 2007). Alternatively, it is tempting to speculate that RS1 could be involved in the PI3K-Akt signalling pathway, which shows increased activity after inhibition of AR in mouse prostate cancer models and xenograft tumours (Carver *et al.*, 2011), especially as Uniprot infers binding to phosphatidylinositol phosphates as one of the predicted molecular functions of RS1, though this has no experimental confirmation (Uniprot identifier: O15537).

The finding that RS1 may be more highly expressed in recurrent patients would agree with the data analysed at the mRNA level (Chapter 3.2.2). No differences were detected between the nuclear staining of RS1 in benign and malignant prostate tissue, or between recurrent and non-recurrent prostate tumours, which would be consistent with the nuclear reaction being a non-specific staining reaction. However, the fact that RS1 appears to be amplified in a subset of metastatic and castration resistant prostate carcinomas makes it an interesting target for further study. Ideally, such a study would include a larger cohort of patients with paired biopsies from diagnosis and after treatment, or perhaps paired castration-resistant or metastatic material, to investigate the prognostic relevance of RS1 expression and its amplification during prostate cancer progression. Given that this was the first time that RS1 expression has been reported outside of the eye, it would also be necessary to verify antibody specificity using positive control tissue on IHC or western blot. If RS1 is expressed in prostate cancer, further studies must also focus on determining its functional roles in prostate cancer cell lines. If RS1 is not the protein recognised by this antibody in prostate cancer tissue, then it may still have merit as a biomarker. For example, the proliferation marker Ki-67 was first identified as being an antigen associated with proliferation and marked by the monoclonal antibody Ki-67, even before it was characterised at the molecular level (Gerdes *et al.*, 1984). However, in this situation, the antigen that anti-RS1 recognises in prostate tumours must also be identified.

4.3.8 Limitations of the Bath Cohort and conclusions

The Bath Cohort data has limited statistical power due to the small sample size. The sample size of this study is insufficient to draw correlations or statistical inferences regarding survival time, node status, metastases and tumour stage, meaning that many important associations with prognostic indicators could have been missed. Indeed, a bias in sampling with higher proportions of recurrent patients to non-recurrent might have impaired the ability of the study to statistically resolve differences. Heterogeneity in patient treatments might also influence the study's ability to resolve differences, as mechanisms of recurrence for each therapy modality will be different. Additionally, while there was no statistical association between any biomarker expression and Gleason grade, none of the patients in this study were assigned a primary Gleason pattern of 5 (Table 4.1), which is the poorest prognostic category in the Gleason grading system. Thus, following up any individual biomarker would require a larger cohort of patients, such as on a tissue microarray, where many patients can be analysed on the same slide for high throughput data acquisition. We have chosen a promising potential biomarker, Syndecan-1, to be analysed in this manner and the results of Bath Cohort and tissue microarray analysis for this potential biomarker are described in the following chapter. BMI1, RS1 and ZSCAN4 would also be suitable candidates for further analysis using this method.

5 SYNDECAN-1 POSITIVE STROMAL CELLS ARE FOUND IN PROSTATE TUMOURS

5.1 Introduction

One potential biomarker for prognostication in prostate cancer is the heparan sulphate proteoglycan Syndecan-1 (SDC1).

5.1.1 Functional roles of SDC1

Otherwise known as CD138, SDC1 is a heparan sulphate transmembrane proteoglycan and is a well-established marker of plasma cells that is expressed in myeloma and in plasma cell lines (Maiga *et al.*, 2014). It is the principle syndecan expressed on epithelial cells (Gharbaran, 2015). As a co-receptor for integrins and growth factors, it plays diverse roles in cell adhesion, migration and cell signalling. In the context of epithelial cells, SDC1 has been shown to act as a co-receptor with α2β1 integrin to facilitate adhesion to type I collagen, and appears to be important for forming focal adhesions in a wide variety of cells including head and neck squamous cell carcinoma cell line HSC-3, Chinese hamster ovary cells and the breast cancer cell line MDA-MB-231 (Ishikawa and Kramer, 2010; Vuoriluoto *et al.*, 2008). The ectodomain of SDC1 is necessary for activation of α2β1 integrin (Altemeier *et al.*, 2012). SDC1 expression

promotes adhesion and cell spreading to collagen substrates, and reduces invasion through them (Vuoriluoto *et al.*, 2011; Ishikawa and Kramer, 2010). This is accompanied by regulation of Rho GTPases, with increased RhoA signalling and decreased Rac signalling occurring as a result of SDC1 binding (Ishikawa and Kramer, 2010). Conversely, downregulation of SDC1 in HSC-3 cells results in decreased focal adhesions, increased lamellipodia formation and cell migration (Ishikawa and Kramer, 2010). In B2b lung carcinoma cells, SDC1 silencing results in a similar increase in cell migration that is accompanied by a decreased focal adhesion lifespan, suggesting that the role of SDC1 in this process is to stabilise focal adhesions (Altemeier *et al.*, 2012). In this cell line the transmembrane domain of SDC1 appears to be necessary for this function and the consequent negative regulation of cell migration (Altemeier *et al.*, 2012).

SDC1 can also couple integrin activation to other receptor signalling events, such as those of the growth factor receptor IGF1R. SDC1 is required for the activation of αvβ3 integrin at focal adhesions in HUVEC cells (Beauvais and Rapraeger, 2010), where ECM-bound SDC1 can cluster at focal adhesions and form a ternary complex together with αvβ3 integrin and IGF1R. This function requires the SDC1 ectodomain (Beauvais and Rapraeger, 2010), and the domain appears to be important for cell-ECM adhesion in mesenchymal cells as well, such as B6FS fibrosarcoma cells and STAV-AB mesothelioma cells (Zong *et al.*, 2011). SDC1 overexpression in mesenchymal cells promotes adhesion and reduces migration and chemotaxis when examined at the functional (Zong *et al.*, 2011) or gene expression levels (Szatmari *et al.*, 2012).

Another function of SDC1 is growth factor binding through the heparan sulphate chains on its ectodomain (Gharbaran, 2015). The inflammatory cytokine Macrophage Migration Inhibitory Factor (MIF) has been found to bind to SDC1 on the surface of lung carcinoma cell line A549 as well as MDA-MB-231 cells, where it stimulates cell migration in a SDC1-dependent fashion (Pasqualon *et al.*, 2016). The ectodomain can also be shed through cleavage by matrix metalloproteinases, creating a soluble form which may have pro-angiogenic roles in tumours, for example in myeloma (Gharbaran, 2015), and the shed ectodomain in itself may possess prognostic utility (Szarvas *et al.*, 2016b). Interestingly, shed SDC1 from myeloma cell lines has been shown to enter the nucleus of other cells, such as bone marrow-derived stromal cells (Stewart *et al.*, 2015) – a process that depends upon binding of growth factors such as HGF to the shed protein, which it also shuttles to the nucleus. Similarly, in STAV-AB mesothelioma

cells, FGF2 and SDC1 colocalise with one another in the nuclear compartment, whereas its receptor FGFR1 does not (Zong *et al.*, 2009). The roles of shed SDC1 in the nucleus are currently poorly understood, but in these stromal cells it binds to p300 and inhibits its histone acetyltransferase activity (Stewart *et al.*, 2015), suggesting that it may be involved in the control of gene transcription. There is also evidence that SDC1 signalling is involved in maintaining the tumour-initiating cell (TIC)/cancer stem cell (CSC) population in breast carcinoma (Alexander *et al.*, 2000; Liu *et al.*, 2004) and prostate carcinoma (Shimada *et al.*, 2013) and can regulate the expression of several microRNAs that suppress expression of stem cell markers such as SOX-2, OCT4 and Nanog (Fujii *et al.*, 2015). Suppression of SDC1 gene expression also reduces the amplification of the putative CSC population in PC-3 cells following treatment with docetaxel (Shimada *et al.*, 2013), suggesting that it could exert effects of relevance to prostate cancer patients undergoing therapy. Overall, it appears that SDC1 primarily functions in growth factor binding, regulation of the actin cytoskeleton, cell adhesion and cell migration in both epithelial cells and mesenchymal cells. However, there may be additional functions of SDC1 that are not yet completely understood, such as regulation of the putative CSC population and stem cell factor expression.

5.1.2 SDC1 expression in the human prostate and its association with cancer

The expression of SDC1 has been linked to prognosis and treatment response in a large range of cancer types, including haematolymphoid malignancies and solid tumours (Gharbaran, 2015; Nguyen *et al.*, 2013; Zellweger *et al.*, 2003; Juuti *et al.*, 2005; Lee *et al.*, 2014; Ramani and Sanderson, 2014), but in prostate cancer the role of SDC1 in prognosis is controversial. It is expressed in the basal epithelial layer of normal prostate glands, and this expression is often lost in prostate carcinoma cells (Chen *et al.*, 2004a; Shimada *et al.*, 2009; Brimo *et al.*, 2010; Shariat *et al.*, 2008; Fujii *et al.*, 2015; Szarvas *et al.*, 2016b).

However, there are reports that hormone-refractory tumours (Brimo *et al.*, 2010) and tumours treated with neo-adjuvant hormone therapy (Shimada *et al.*, 2009) have increased SDC1 expression, suggesting a relationship between expression and androgen independence. There are also reports that expression of SDC1 is associated with high Gleason grade disease and/or early disease recurrence in prostate cancer patients (Chen *et al.*, 2004a; Shimada *et al.*, 2009; Brimo *et al.*, 2010; Shariat *et al.*, 2008; Fujii *et al.*,

2015). Interestingly, an alteration in protein localization, from membranous to cytoplasmic, was found to be associated with a more rapid biochemical recurrence. This indicates that alterations in SDC1 expression and localization could be prognostically important and change throughout the course of the disease. Conversely, others have found that tissue SDC1 expression is not associated with recurrence-free survival (Purushothaman *et al.*, 2010; Brimo *et al.*, 2010) or Gleason grade in prostate carcinoma (Brimo *et al.*, 2010), and even concluded that lower levels of SDC1 expression were associated with a high Gleason grade (Gharbaran, 2015). Therefore, the potential utility of SDC1 as a biomarker for prostate cancer is far from clear.

Taken together, there is accumulating evidence that SDC1 is an important multifunctional molecule which shows altered expression in prostate cancer, but its potential role as a biomarker for prostate cancer remains controversial. Given the evidence concerning SDC1 expression in prostate tumours, I predict that SDC1 expression will be predominantly localised to basal cells in benign prostate tissue, and that this expression will be lost in the majority of tumours. Additionally, I predict that SDC1 expression in prostate carcinoma epithelium will be positively associated with biochemical recurrence.

5.1.3 Aims

Given the diverse functional roles played by SDC1 and the conflicting evidence surrounding its possible utility as a biomarker, we aimed to investigate SDC1 expression further by IHC in three independent cohorts of prostate carcinoma samples. This chapter will:

i. Assess epithelial expression patterns of SDC1 in benign and malignant prostate tissue, and the ability of this biomarker to distinguish between normal and disease states.

ii. Identify any relationships between epithelial SDC1 expression patterns and biochemical recurrence status after 5 years of retrospective follow-up.

iii. Further characterise the cell types marked by SDC1 expression in prostate tumours using IHC and IF techniques.

5.2 Methods

5.2.1 Patient cohorts in this chapter

Data obtained for this study came from three separate cohorts of patients with associated formalin-fixed, paraffin-embedded (FFPE) prostate material available. A subset of the Bath Cohort (detailed in Chapter 4) was used for preliminary staining and analysis of cell types by fluorescent co-staining, and consisted of six prostate cancer patients with confirmed SDC1+ stromal cells as assessed by anti-SDC1 IHC. In addition, one block of adjacent normal prostate tissue was used as a positive control for SDC1 staining. These samples were obtained retrospectively from patients treated for prostate cancer at the Royal United Hospital, Bath, United Kingdom, between 1997 and 2008. The samples were obtained following ethical approval from the NHS Ethics Committee in 2013 (REC reference: 13/WS/0153; IRAS project ID: 112241). Informed written consent was obtained from all patients. NHS ethics approval also covered the other two patient cohorts. These were used for anti-SDC1 immunohistochemistry and were obtained from US Biomax (Rockville, MD, USA). The first cohort (PR1921) consisted of 80 cases of prostate adenocarcinoma with 8 cases of adjacent normal prostate tissue and 8 cases of histologically normal prostate tissue, with duplicate 1mm cores per cancer case for a total of 196 cores. The second cohort (PR803b) consisted of 71 cases of prostate adenocarcinoma, 2 cases of prostate leiomyosarcoma, 1 case of benign prostate hyperplasia and 6 cores of histologically normal prostate, with a single 1.5mm core per case. Each cohort was contained on a single tissue microarray slide. A summary of the patient characteristics of tissue microarray cohorts is given in Table 5.1. Immunohistochemistry was performed as described in Chapter 2.2.2. All anti-SDC1 immunohistochemistry was performed using the mouse polyclonal B-A38, unless otherwise stated.

Table 5.1: Clinical characteristics of tissue microarray cohort 1 (PR1921) and cohort 2 (PR803B).

	Cohort 1 (PR1921)	%	Cohort 2 (PR803b)	%
Number of Patients	80		71	
Age:				
Range	20-85		20-87	
Mean	67.85		67.6	
Median	69.5		70	
<50	2	2.5	3	4.2
50-59	5	6.3	4	5.6
60-69	33	41.3	27	38.0
>=70	40	50.0	37	52.1
Clinical Stage:				
I	6	7.5	2	2.8
II	37	46.3	33	46.5
III	14	17.5	10	14.1
IV	22	27.5	26	36.6
ND	1	1.3	0	0.0
Primary Gleason Grade:				
3	8	10.0	19	26.8
4	36	45.0	24	33.8
5	30	37.5	24	33.8
ND	6	7.5	4	5.6
Overall Gleason Grade:				
<=Gleason 3+4	8	10.0	16	22.5
>=Gleason 4+3	66	82.5	51	71.8
ND	6	7.5	4	5.6
T Category:				
T1	4	5.0	2	2.8
T2	47	58.8	40	56.3
T3	22	27.5	21	29.6
T4	6	7.5	8	11.3
ND	1	1.3	0	0.0
N Category:				
N0	64	80.0	51	71.8
N1	15	18.8	18	25.4
N2		0.0	2	2.8
ND	1	1.3	0	0.0
M Category:				
M0	64	80.0	49	69.0
M1	15	18.8	22	31.0
ND	1	1.3	0	0.0

5.2.2 Multiplex IF on FFPE Tissue

FFPE tissue blocks were chilled to 4ºC and sections were cut at 5µm thickness, floated out onto silane-coated slides and air-dried as before. Prior to staining, sections were deparaffinised and rehydrated as described previously. Washing, permeablisation and antigen retrieval was all performed as described previously. Sections were ringed with an ImmEdge hydrophobic barrier pen (Vector Laboratories) and blocked for 30 minutes using IHC blocking buffer as in IHC, before incubation with primary antibodies (Table 2.1) diluted in DAKO Antibody Diluent (DAKO) overnight at 4ºC in a humidified chamber. For double immunolabelling experiments, two antibodies recognising different antigens and raised in different host species were mixed together at their respective dilutions for simultaneous labelling.

The next day, sections were washed three times in PBST and fluorescent labelling was performed in the dark for 30 minutes at room temperature with gentle agitation. Alexa Fluor 488-labelled goat anti-mouse IgG and Alexa Fluor 568-labelled goat anti-rabbit IgG antibodies were used, and both fluorescently labelled antibodies (Table 2.2) were mixed together at 1:200 dilutions in DAKO Antibody Diluent (DAKO) for simultaneous labelling. After incubation, sections were washed three times in PBST and then a modified protocol from Viegas et al. (2007) was used to quench endogenous tissue autofluorescence: samples were incubated with a filtered solution of 0.1% (w/v) Sudan Black B in 70% (v/v) ethanol for 10 minutes to quench endogenous tissue autofluorescence. Sections were rinsed briefly with PBST to remove any stain debris, then washed three times for 5 minutes each before mounting in Mowiol with 600nM DAPI (Sigma-Aldrich) for nuclear counterstaining. Samples were allowed to dry overnight in the dark at room temperature, or for at least two hours at 37ºC, before imaging on an LSM510 META laser scanning confocal microscope. Imaging was conducted using 405nm, 488nm and 543nm laser lines for excitation of fluorescent dyes and fields of view were collected under a 40x objective. Samples were stored at 4ºC in the dark and imaged within several days of staining for best results.

5.2.3 SDC1 Antibodies

Two SDC1 antibodies are used in this thesis chapter. The epitope for mouse monoclonal anti-SDC1 antibody clone B-A38 is unidentified, but the original immunogen was the SDC1-expressing plasmacytoma cell line U266, and it has been shown to robustly mark a CD45-expressing plasma cell population by flow cytometry (Maiga *et al.*, 2014), a

cell population which has long been known to express SDC1 (Chilosi *et al.*, 1999). The rabbit anti-SDC1 antibody clone H-174 was generated by immunization with amino acids 82-256 of human SDC1, corresponding to the human SDC1 ectodomain. Shed SDC1 ectodomain immunoprecipitated with this antibody has been successfully identified by western blot with another mouse monoclonal antibody to the protein (Schmidt *et al.*, 2005), suggesting that the antibody is specific to SDC1 and can detect shed SDC1.

5.2.4 Scoring

5.2.4.1 Scoring of SDC1 IHC data

For assessment of IHC staining on the Bath Cohort and the two tissue microarrays, whole cores were examined under a 10x objective to determine the extent and intensity of glandular cytoplasmic/membranous staining of SDC1. This was assessed on a scale of 0-3 for intensity: 0 – no staining; 1 – Weak staining; 2 – Moderate staining; 3 – Strong staining. Extent was scored on a scale based on the proportion of glandular cells stained: 0-4: 0 – no staining; 1 - <25% cells stained; 2 – 25-50% cells stained; 3 – 50-75% cells stained; 4 – 75-100% cells stained. The score for SDC1 glandular staining was determined as the sum of extent and intensity scores, allowing for a maximum score of 7.

For quantification of SDC1+ stromal cell burden, fields of view amounting to all SDC1+ stromal cells in every core were manually acquired under a 20x objective. Counting and measuring of the resulting images was performed using ImageJ v2.0.0. Each stromal cell was identified manually and the following size and shape descriptors were measured: area, circularity, aspect ratio, roundness and solidity. Mean values were taken for each measure within each cohort as a whole, and for each patient.

5.2.4.2 Scoring of SDC1 IF co-staining data

To score SDC1+ stromal cells for their colocalisation with test markers, images from each experimental run were pooled, and each SDC1+ cell was manually scored for positivity of the co-stained marker. The percentage of marker-positive cells was calculated as the percentage of SDC1+/Marker+ cells relative to the total number of SDC1+ cells in an experimental run. Data was plotted in R version 3.2.0.

5.2.5 Data Analysis

To analyse the distribution of SDC1+ cell burden amongst different patient groups in both TMA cohorts, we classified patients into SDC1 stromal cell positive and negative categories based on whether SDC1+ stromal cells were present or absent in the core. In the case of PR1921, where duplicate cores were scored, patients were classified as positive if at least one core contained SDC1+ stromal cells. The distributions of SDC1+ and SDC1- patients amongst different levels of clinical variables were tested with the chi-squared test performed in R version 3.2.0 with default parameters. Clinical variables assessed were: age; stage; Gleason grade; T, N and M stages; Presence of SDC1 staining in the glandular compartment was also scored as positive or negative, and tested for association as with clinical variables. A p-value of less than 0.05 was considered statistically significant. Graphs were also generated in R version 3.2.0.

5.3 Results

5.3.1 SDC1 expression patterns in benign and malignant prostate tissue

To begin to investigate the expression of SDC1, I subjected the Bath Cohort, including benign and malignant prostate tissue, to anti-SDC1 IHC. Expression was found in the cell membrane and cytoplasm of the prostate epithelium basal cell layer in histologically normal tissue (Figure 5.1A). In prostate carcinoma tissues there was either a loss of epithelial staining or a widespread expression of the proteoglycan throughout the cytoplasm and/or cell membranes of prostate tumour epithelium (Figure 5.1A). Both normal and tumour staining patterns are consistent with previous studies (Zellweger *et al.*, 2003; Chen *et al.*, 2004a; Shariat *et al.*, 2008; Ledezma *et al.*, 2011; Contreras *et al.*, 2010; Szarvas *et al.*, 2016b). Interestingly, in addition to these staining patterns, a population of scattered SDC1-positive cells were found in the stroma in a subset of cancer cases (Figure 5.1A, arrows). Since these SDC1-positive stromal cells had not been demonstrated before, subsequent experiments focused on validating and characterising them, as well as assessing their relevance as prognostic biomarkers.

The prevalence of SDC1-positive stromal cells was further investigated in two additional cohorts of 80 and 71 prostate cancer patients, respectively. Each of the two additional cohorts was represented by a single tissue microarray, and tumour tissue was accompanied by histologically normal prostate tissue (in both cohorts) and 8 cases of matched adjacent normal tissue (in cohort 1). The scattered SDC1-positive stromal cells

were found in a subset of tumour samples from each of the two additional cohorts, either proximal or distal to an epithelial region. A range of examples from all three cohorts are shown in Figure 5.1A and B. SDC1 staining was found to be localized to the membranes and/or cytoplasmic regions of the SDC1-positive cells, and often appeared in close proximity to areas of inflammation which were characterised by the stromal infiltration of cells with compact, rounded and hyperchromatic nuclei that are thought to be lymphocytes (Figure 5.1B, arrowheads).

To confirm that the staining was due to the presence of SDC1, and not cross reactivity from the mouse antibody, a second anti-SDC1 antibody raised in rabbit was used for IHC and IF staining on the Bath Cohort. IHC staining of sequential sections with the two different anti-SDC1 antibodies was performed. The rabbit anti-human SDC1 antibody clone H-174 demonstrated a higher intensity of glandular cytoplasmic staining than the mouse anti-human SDC1 antibody B-A38, as shown by staining of consecutive sections in IHC (Figure 5.2A), but both antibodies stained PCSP cells (Figure 5.2A, arrows). It is possible that the increased glandular cytoplasmic staining of H-174 is due to its recognition of other domains of the protein, such as the soluble ectodomain in addition to the full length SDC1 protein (Schmidt *et al.*, 2005). Double IF staining using both antibodies showed colocalisation in SDC1-positive stromal cells (Figure 5.2B, arrows), as well as demonstrating equivalent labelling of the basal cell layer in benign prostate tissue (Figure 5.2B).

In summary, in addition to previously documented epithelial staining patterns of SDC1, a subset of prostate cancer samples contained a population of scattered SDC1 positive cells that can be detected using two different anti-SDC1 antibodies and in 3 patient cohorts. To our knowledge, this cell type has not been previously described. We call these cells Prostate Cancer SDC1 Positive (PCSP) cells.

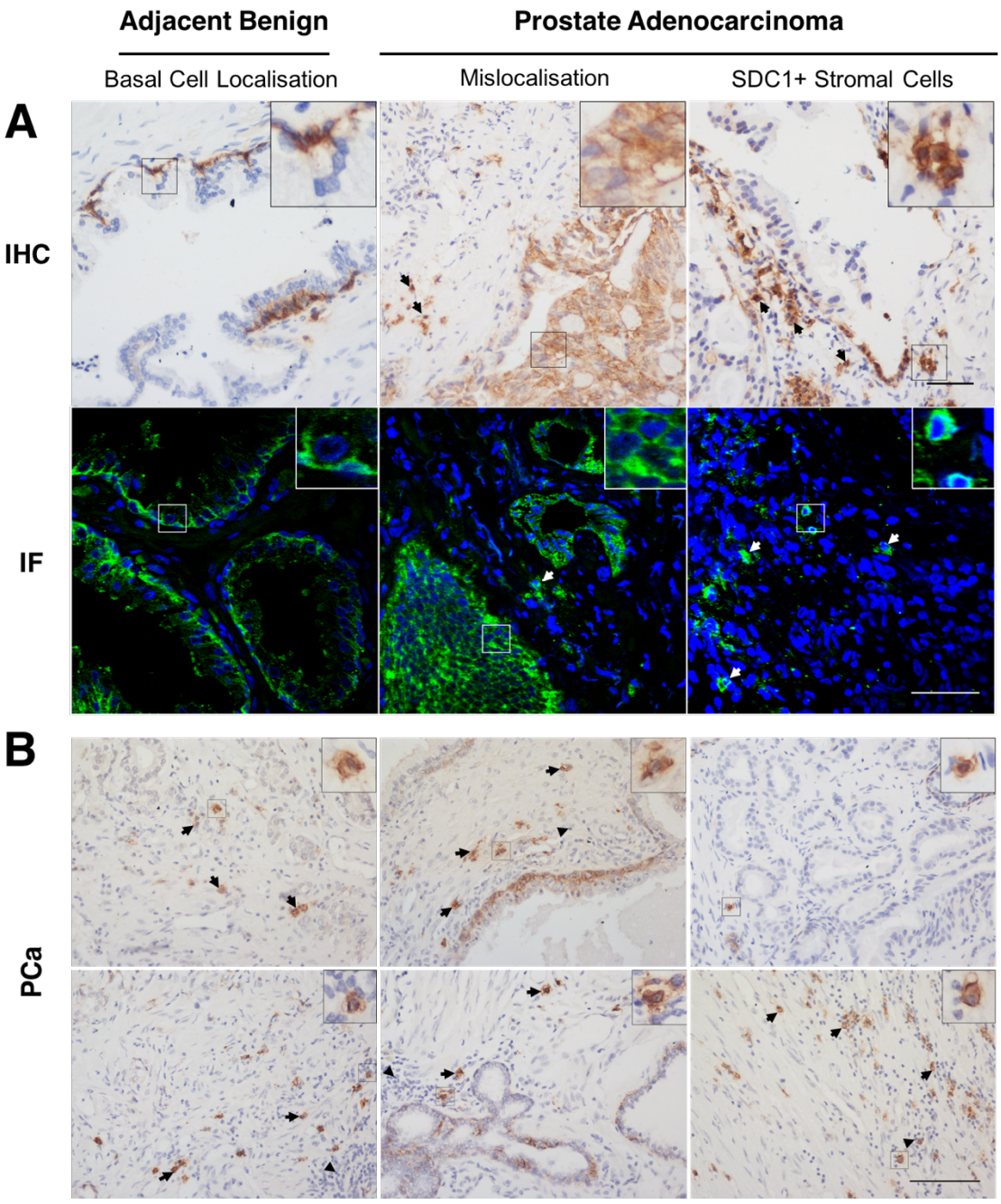

Adjacent Benign
Prostate Adenocarcinoma
Basal Cell Localisation
Mislocalisation
SDC1+ Stromal Cells
A
IHC
IF
B
PCa

Figure 5.1: SDC1 expression is altered in prostate carcinoma with the appearance of Prostate Cancer SDC1-Positive Stromal (PCSP) cells in a subset of cases.

(A) In a subset of prostate cancer cases, PCSP cells are found scattered in the stroma (shown inset and by arrows). While not found in histologically normal tissue, PCSP cells can be found in the contexts of both epithelial SDC1 alterations in tumours, including expression across all epithelial cells, and where SDC1 expression is lost in the epithelium entirely. These cells are demonstrable by both IHC and IF staining techniques in FFPE prostate tissue. Representative images of samples from the Bath Cohort are shown. (B) Examples of PCSP cells detected by anti-SDC1 IHC with the mouse antibody clone B-A38. All PCSP cells are located in the stroma and tend to appear near regions containing putative lymphocytes with compact, round, hyperchromatic nuclei (arrowheads). Representative images of samples from TMA cohorts 1 and 2 are shown. Nuclei are counterstained with haematoxylin (blue) in IHC, or with DAPI (blue) in IF. IHC scale bars – 100um; IF scale bars – 50um.

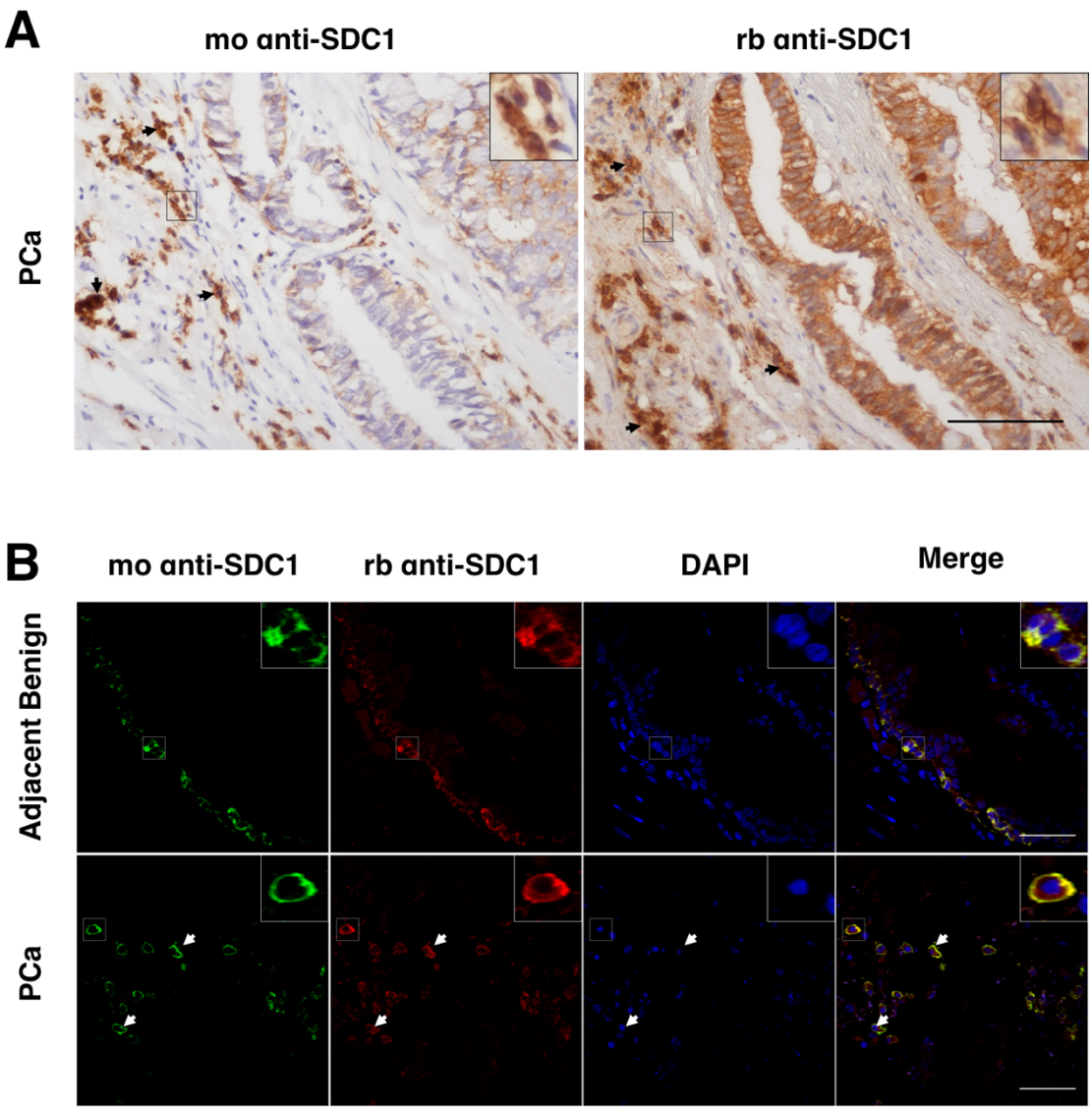


Figure 5.2: The presence of PCSP cells is confirmed by staining with two distinct anti-SDC1 antibodies.

(A) IHC staining of adjacent tissue sections of a representative prostate tumour (PCa) with a mouse (mo) anti-SDC1 antibody (left) and a rabbit (rb) anti-SDC1 antibody (right), showing increased cytoplasmic staining with the rabbit antibody, but detection of PCSP cells (arrows) by both antibodies. (B) The specificity of the mouse antibody (green) is further confirmed by IF co-staining with another SDC1-specific antibody raised in rabbit (red), showing colocalisation in both benign (basal cells, inset) and tumour tissues (PCSP cells, inset and arrows). Nuclei are counterstained with haematoxylin (blue) in IHC, or with DAPI (blue) in IF. IHC scale bars – 100um; IF scale bars – 50um.

5.3.2 Epithelial cytoplasmic SDC1 expression was reduced and PCSP cell burden was elevated in malignant prostate stroma of the Bath Cohort

To quantify SDC1 expression patterns in both benign and malignant prostate tissue, the cytoplasmic scoring system was used to quantify epithelial staining, and the number and area of PCSP cells was quantified and expressed as the total number in five fields and the total area per field of view, expressed as an average of five fields respectively. There was a trend towards a reduced cytoplasmic score in malignant prostate when compared to benign prostate tissue (Figure 5.3A), but this difference was not significant ($p=0.6689$). A subset of malignant patient samples demonstrated a high burden of PCSP cells (Figure 5.3B), with 12/28 (42.9%) having a mean burden of greater than 3 PCSP cells per field compared with 1/5 (20%) of benign tissues. However, the sample size was insufficient to demonstrate a significant difference in PCSP cell burden between the benign and malignant prostate tissue ($p=0.2726$). Similarly, when measuring the total area of PCSP cells in patient tissues, there was no significant difference ($p=0.6643$) although a trend towards increased area in malignant tissues was apparent (Figure 5.3C).

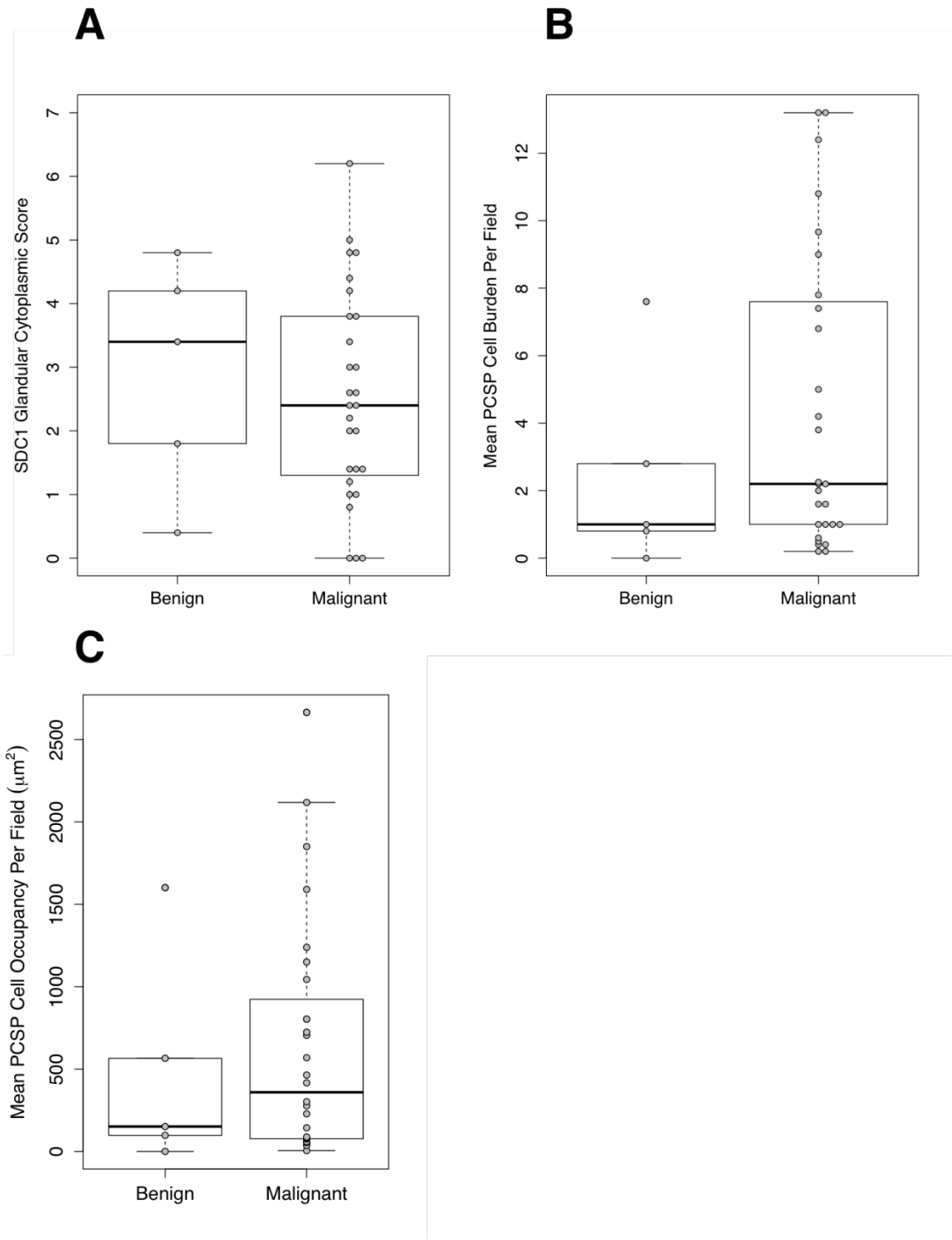

A
SDC1 Glandular Cytoplasmic Score
0
1
2
3
4
5
6
7
Benign
Malignant
B
Mean PCSP Cell Burden Per Field
0
2
4
6
8
10
12
Benign
Malignant
C
Mean PCSP Cell Occupancy Per Field (μm²)
0
500
1000
1500
2000
2500
Benign
Malignant

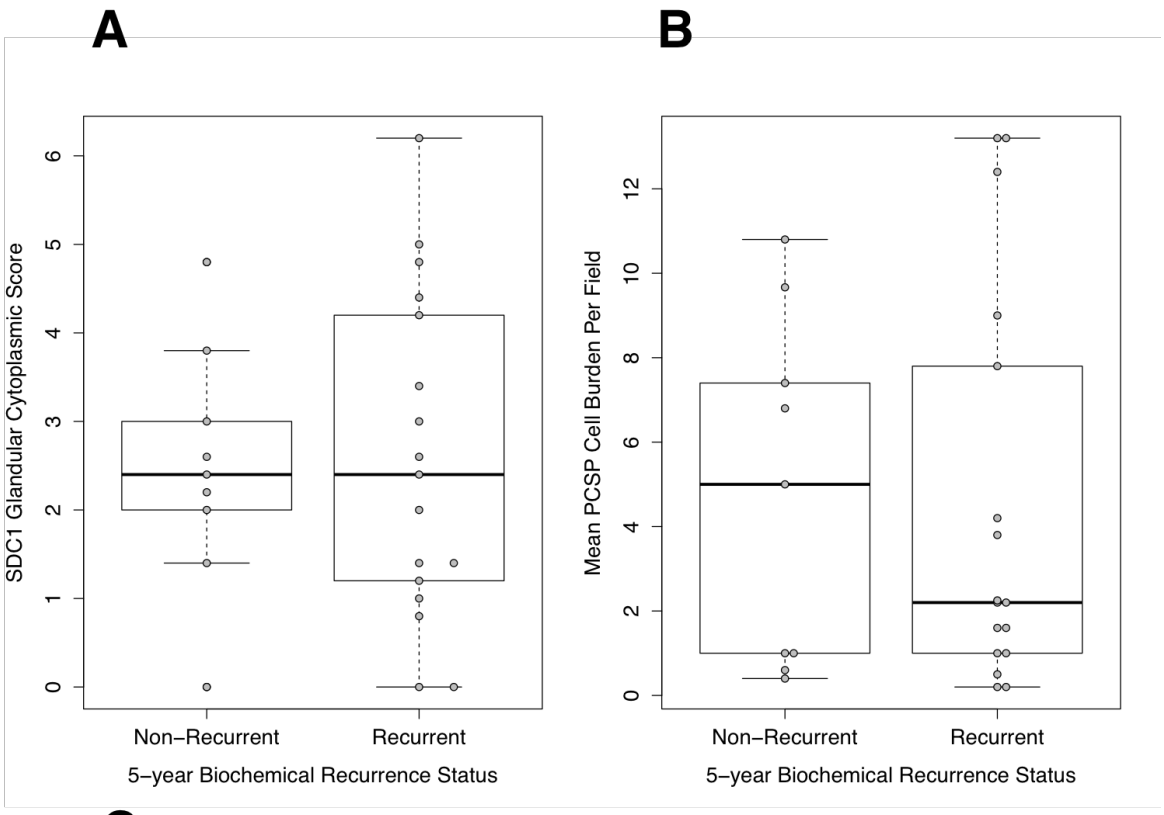
Figure 5.3: Quantification of cytoplasmic SDC1 staining and Prostate Cancer SDC1 Positive Stromal (PCSP) cells in benign and malignant prostate tissue.

IHC staining of SDC1 was quantified in the Bath cohort using the cytoplasmic score for prostate epithelium, and the number and area of PCSP stromal cells was measured. SDC1 staining was compared between benign and malignant prostate epithelium: (A) glandular cytoplasmic score showing a slight trend towards reduced cytoplasmic staining of malignant prostate epithelium ($p=0.6689$); (B) mean frequency of PCSP cells per 20x field of view, which appeared to increase in malignant tissue, although not statistically significant ($p=0.2726$), and; (C) mean total area of PCSP cell occupancy per 20x field of view, which showed a similar trend to cell burden, and was not statistically significant ($p=0.6643$). Two-sample t-tests were conducted to determine statistical significance for each set of conditions. n=5 benign samples, 28 malignant samples.

5.3.3 Associations between SDC1 staining and outcome in the Bath Cohort

In order to determine the relevance of all forms of SDC1 staining patterns in a prognostic biomarker setting, we dichotomised all prostate cancer patients with available tissue for SDC1 staining in the Bath Cohort, based on 5-year biochemical recurrence-free survival status (n=26 patients, 9 non-recurrent and 17 recurrent). There was no significant difference between SDC1 glandular cytoplasmic staining scores in 5-year recurrence-free and non-recurrent prostate tumour tissue (Figure 5.4A) (p=0.775). PCSP cell burden was similar between recurrent and non-recurrent patient samples, and there was no significant difference (Figure 5.4B) (p=0.952). However, when PCSP cells occupying the tissue were measured, there was a trend towards reduced total area of occupying cells in recurrent patient samples compared to non-recurrent samples (Figure 5.4C), although it was not statistically significant (p=0.1597). Any conclusions on this relationship would require analysis of a larger cohort of patients with available recurrence data.

To summarise findings from the Bath Cohort, there were trends towards an increased PCSP cell burden in malignant tissue compared to benign tissue, and reduced burden in recurrent prostate tumours compared to non-recurrent tumours, but the small cohort was unable to demonstrate statistical significance in these findings. These trends in a previously uncharacterised cell type led us to verify our observations in larger datasets. To do this, we expanded our analysis of PCSP cell burden to look at the two tissue microarray cohorts stained for SDC1.


A
SDC1 Glandular Cytoplasmic Score
0
1
2
3
4
5
6
Non−Recurrent
Recurrent
5−year Biochemical Recurrence Status
B
Mean PCSP Cell Burden Per Field
0
2
4
6
8
10
12
Non−Recurrent
Recurrent
5−year Biochemical Recurrence Status

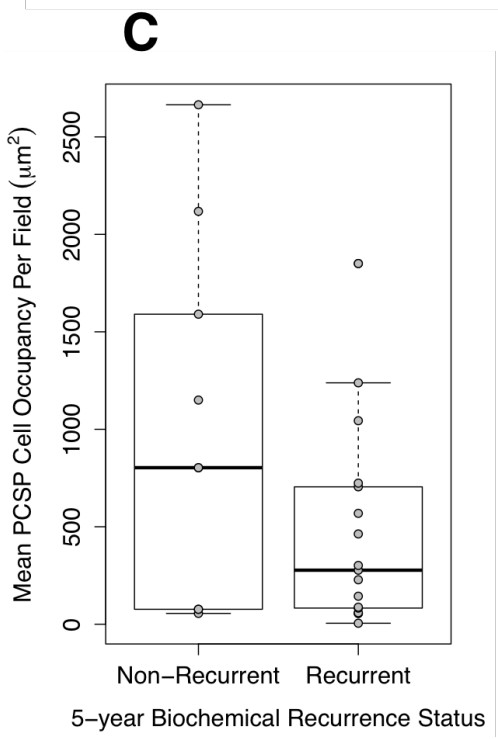

C
Mean PCSP Cell Occupancy Per Field (μm²)
0
500
1000
1500
2000
2500
Non−Recurrent
Recurrent
5−year Biochemical Recurrence Status

Figure 5.4: Quantification of epithelial cytoplasmic SDC1 staining and PCSP cell burden and total area in 5-year recurrence-free and recurrent prostate tumours from the Bath Cohort.

IHC staining of SDC1 was quantified in the Bath cohort using the cytoplasmic score for prostate epithelium, and the number and area of PCSP stromal cells was measured. SDC1 staining was compared between recurrent and non-recurrent prostate tumour samples: (A) glandular cytoplasmic score showing similar levels of staining between recurrent and non-recurrent samples (p=0.775); (B) mean frequency of PCSP cell occurrence per 20x field of view over five fields per sample, which appeared to be similar between both groups (p=0.952), and; (C) mean total area of PCSP cell occupancy per 20x field of view, which showed a trend towards decreased occupancy in recurrent samples, although this did not reach significance (p=0.1597). Two-sample t-tests were conducted to determine statistical significance for each set of conditions. *N*=9 non-recurrent samples, 17 recurrent samples.

5.3.4 Analysis of PCSP cells in patient tissue microarrays reveals Gleason pattern 5 tumours with a particularly high burden

Using the clinical metadata available for both prostate tissue microarray cohorts, we wished to determine if there was a relationship between SDC1+ stromal cell burden and clinicopathological features. These included age, Gleason grade and TNM stages of all patients in the study. In the first cohort, cores of histologically normal tissue adjacent to prostate cancer were included, as well as non-diseased normal prostate control cores. Patient characteristics for both cohorts are summarized in Table 5.1, and have similar patient distributions among the clinical variables.

PCSP cells were found in 28 of 80 cases (35%) of patients with prostate cancer in cohort 1, but they were not observed in any of the 8 control cases of non-diseased patient tissue (Figure 5.5A). 7 of 28 positive cancer cases (25%) were PCSP-positive in both cores. In cohort 2, PCSP cells were found in 15 of 71 prostate cancer cases (21.1%). In contrast with both TMA cohorts, the Bath cohort displayed some evidence of PCSP cell burden in all cases, even if the burden was very low (Figure 5.5B). Interestingly, in cohort 1, PCSP cells were also found in 4 of 8 cases of normal patient tissue that was adjacent to carcinoma (Figure 5.5A), but all of these cases only had a single PCSP-positive core. The presence of PCSP cells in normal adjacent tissue might be explained by being a migratory and/or ubiquitous cell type – a hypothesis that will be investigated further in Chapter 5.3.5.

Chi-squared tests were conducted to determine the association between PCSP cell presence and available clinical parameters for the tissue microarray cases, and the full results are given for cohort 1 (Table 5.2) and cohort 2 (Table 5.3). With respect to Gleason grade, no patient in the Bath cohort presented with a primary Gleason pattern 5 tumour, precluding analysis of all Gleason patterns. However, there was no significant difference in PCSP cell burden between primary Gleason pattern 3 and 4 tumours in the Bath Cohort ($p=0.673$). Interestingly, in both tissue microarray cohorts a subset of Gleason grade five tumours had a particularly high burden of PCSP cells (Figure 5.5C and D). Among cohort 2, there was also a statistically significant association between PCSP cell burden and higher primary Gleason grade ($p=0.0167$). However, in cohort 1 it fell below significance ($p=0.706$).

Similarly, there was a statistically significant association between PCSP cell presence and a lower clinical stage in cohort 1 ($p=0.030$), but this pattern did not hold true in cohort 2 ($p=0.614$). Given that there were only 6 stage I patients in cohort 1, 5 of which

had PCSP cells, it is likely that the under-sampling of the lowest grade tumours contributed to this result in cohort 1. Further studies are needed to confirm if there is a statistically significant and relevant relationship between PCSP burden and primary Gleason grade or clinical stage. In summary, analysis of PCSPs in the tissue microarrays showed that PCSP cells were not observed in the normal tissue analysed, but were found in adjacent normal and in roughly 1/3 of prostate carcinoma tissues, with a subset of high Gleason grade tumours having a particularly high burden. The tissue microarray cohort data supports the findings of the Bath Cohort with respect to PCSP cell burden in benign and malignant prostate epithelium, and additionally indicates that PCSP cells may be found in adjacent normal tissue of prostate tumours.

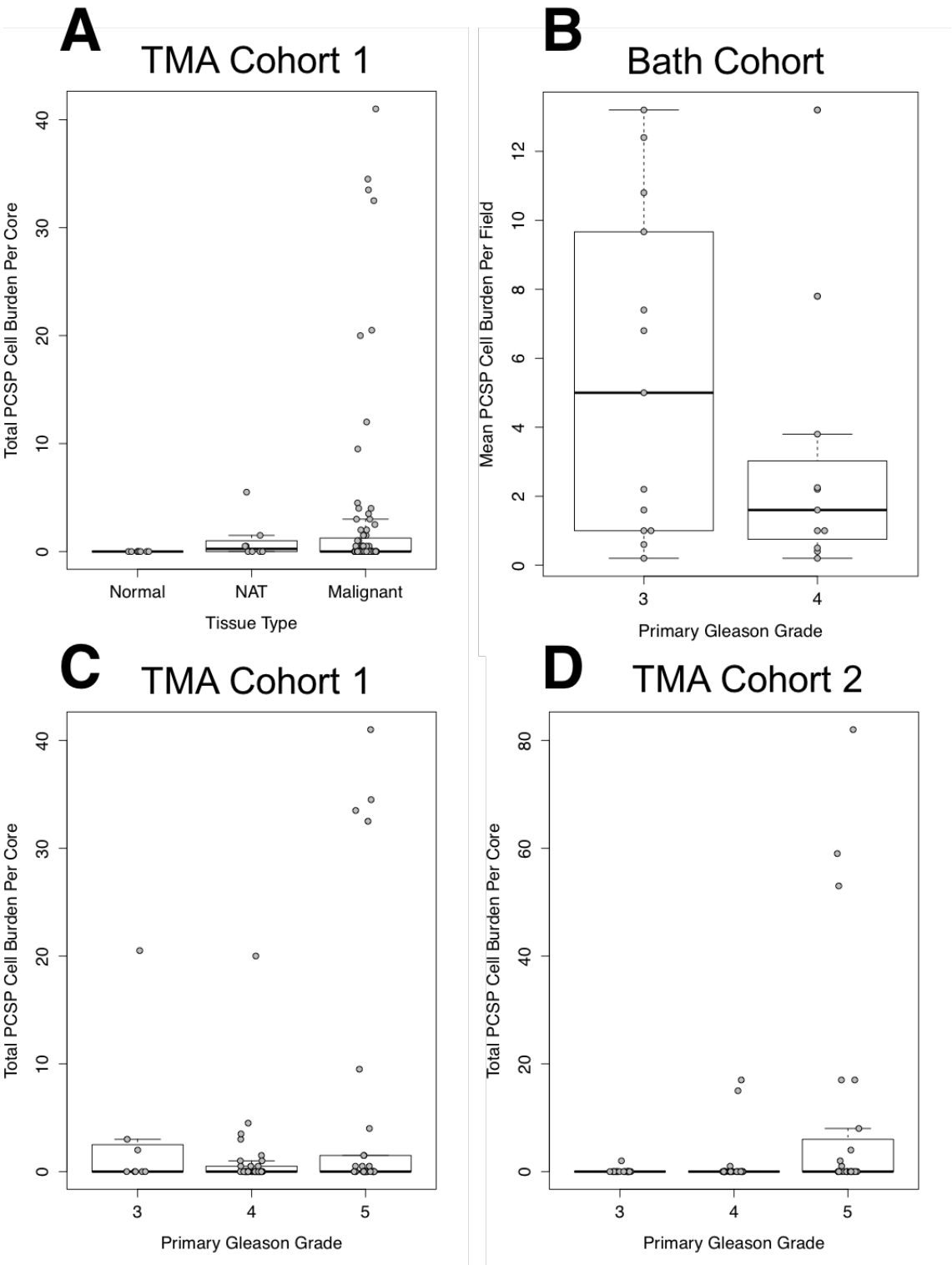

A
TMA Cohort 1
Total PCSP Cell Burden Per Core
0
10
20
30
40
Normal
NAT
Malignant
Tissue Type
B
Bath Cohort
Mean PCSP Cell Burden Per Field
0
2
4
6
8
10
12
3
4
Primary Gleason Grade
C
TMA Cohort 1
Total PCSP Cell Burden Per Core
0
10
20
30
40
3
4
5
Primary Gleason Grade
D
TMA Cohort 2
Total PCSP Cell Burden Per Core
0
20
40
60
80
3
4
5
Primary Gleason Grade

Figure 5.5: PCSP cells are found in normal tissue adjacent to carcinoma in addition to carcinoma tissue, and a subset of Gleason 5 cases have a particularly high burden of PCSP cells.

(A) Plots of PCSP cell burden in cohort 1. PCSP cells are detected in 4/8 normal adjacent (NAT) samples and in 28/80 malignant samples, but not in non-diseased normal samples (0/8). This distribution does not quite reach statistical significance by the chi squared test (p=0.0781). (B) Plots of PCSP cell burden in the Bath Cohort showing burden in both Gleason 3 and 4 cases with no significant difference between groupings (p=0.673). The Bath Cohort does not contain patients with a primary Gleason pattern 5. (C) Plots of PCSP cell burden in cohort 1 and (D) cohort 2 malignant tissues show primary Gleason pattern 5 patients with a particularly high burden in both cohorts, which was statistically significant in cohort 2 (p=0.0167) but not in cohort 1 (p=0.706). The entire core for each specimen was quantified. Tumour cases with available Gleason grading information: Bath Cohort: N=24 tumours; TMA Cohort 1: N=74 tumours; TMA Cohort 2: N=67 tumours.

Table 5.2: Characteristics of patients from Cohort 1 (PR1921) and chi-squared tests between patients with (PCSP+) and without (PCSP-) a PCSP cell burden.

	All Patients	*%*	*PCSP+*	*%*	*PCSP -*	*%*	*P-Value (χ^2 test)*
#Patients	80	100	28	35	52	65	
Age:	Range: (20-85), Mean: (67.85), Median: (69.5)						
<50	2	2.5	0	0.0	2	2.5	
50-59	5	6.3	0	0.0	5	6.3	
60-69	33	41.3	13	16.3	20	25.0	
>=70	40	50.0	15	18.8	25	31.3	0.245
Clinical Stage:							
I	6	7.5	5	6.3	1	1.3	
II	37	46.3	13	16.3	24	30.0	
III	14	17.5	5	6.3	9	11.3	
IV	22	27.5	4	5.0	18	22.5	
ND	1	1.3	1	1.3	0	0.0	0.030
Tumour Grade:							
2	8	10.0	3	3.8	5	6.3	
2--3	30	37.5	9	11.3	21	26.3	
3	36	45.0	12	15.0	24	30.0	
NA	6	7.5	4	5.0	2	2.5	0.910
Primary Gleason:							
3	8	10.0	3	3.8	5	6.3	
4	36	45.0	10	12.5	26	32.5	
5	30	37.5	11	13.8	19	23.8	

NA	6	7.5	4	5.0	2	2.5	0.706
Overall Gleason:							
<=Gleason 3+4	8	10.0	3	3.8	5	6.3	
>=Gleason 4+3	66	82.5	21	26.3	45	56.3	
NA	6	7.5	4	5.0	2	2.5	1.000
T Category:							
pT1	4	5.0	3	3.8	1	1.3	
pT2	47	58.8	16	20.0	31	38.8	
pT3	22	27.5	7	8.8	15	18.8	
pT4	6	7.5	1	1.3	5	6.3	
NA	1	1.3	1	1.3	0	0.0	0.280
N Category:							
N0	64	80.0	23	28.8	41	51.3	
N1	15	18.8	4	5.0	11	13.8	
NA	1	1.3	1	1.3	0	0.0	0.705
M Category:							
M0	64	80.0	25	31.3	39	48.8	
M1	15	18.8	2	2.5	13	16.3	
NA	1	1.3	1	1.3	0	0.0	0.112
SDC1 Epithelial Reactivity:							
Absent	66	82.5	17	21.3	49	61.3	
Present	14	17.5	11	13.8	3	3.8	0.001

Table 5.3: Characteristics of patients from Cohort 2 (PR803b) and chi-squared tests between patients with (PCSP+) and without (PCSP-) a PCSP cell burden.

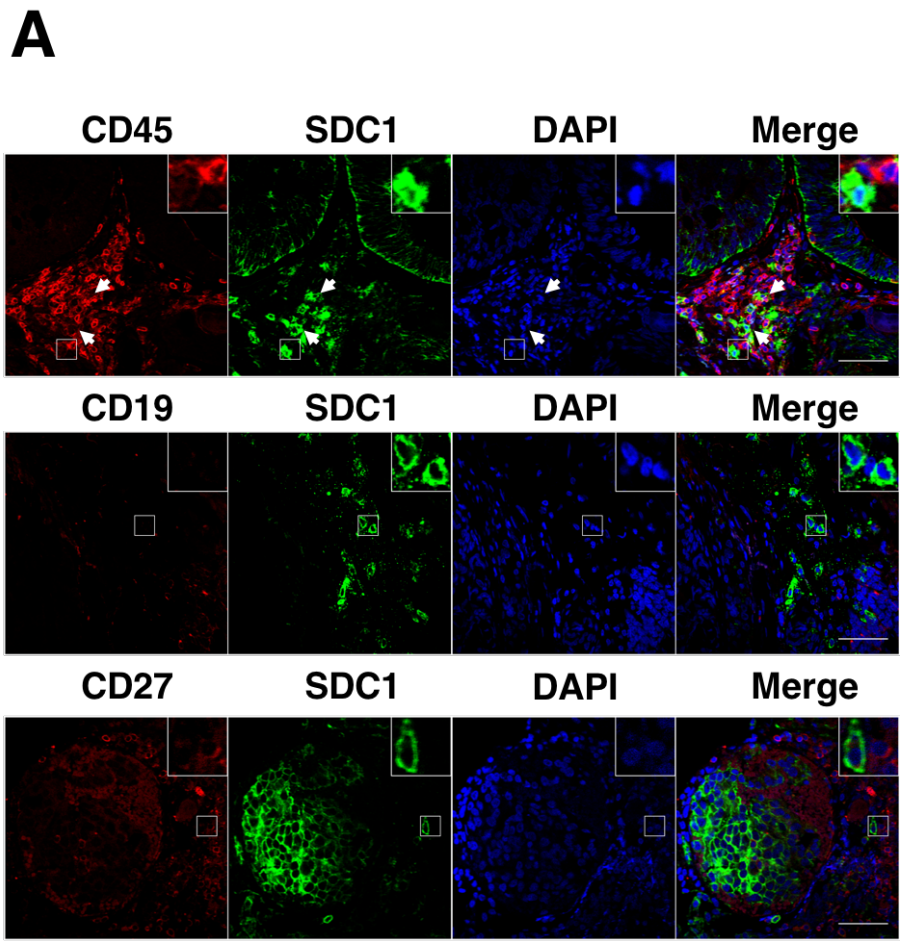
	All Patients	*%*	*PCSP+*	*%*	*PCSP-*	*%*	*P-Value* (χ^2 test)
# Patients	71	100	15	21.1	56	78.9	
Age:	Range: (20-87), Mean: (67.6), Median: (70)						
<50	3	4.2	2	2.8	1	1.4	
50-59	4	5.6	1	1.4	3	4.2	
60-69	27	38.0	5	7.0	22	31.0	
>=70	37	52.1	7	9.9	30	42.3	0.263
Clinical Stage:							
I	2	2.8	1	1.4	1	1.4	
II	33	46.5	7	9.9	26	36.6	
III	10	14.1	1	1.4	9	12.7	
IV	26	36.6	6	8.5	20	28.2	0.614
Tumour Grade:							
2	21	29.6	2	2.8	19	26.8	
3	47	66.2	11	15.5	36	50.7	
NA	3	4.2	2	2.8	1	1.4	1.000
Primary Gleason:							
3	19	26.8	1	1.4	18	25.4	
4	24	33.8	3	4.2	21	29.6	
5	24	33.8	9	12.7	15	21.1	
NA	4	5.6	2	2.8	2	2.8	**0.017**
Overall Gleason:							
<=Gleason 3+4	16	22.5	0	0.0	16	22.5	
>=Gleason 4+3	51	71.8	13	18.3	38	53.5	
NA	4	5.6	2	2.8	2	2.8	0.059
T Category:							
T1	2	2.8	1	1.4	1	1.4	
T2	40	56.3	9	12.7	31	43.7	
T3	21	29.6	3	4.2	18	25.4	
T4	8	11.3	2	2.8	6	8.5	0.635
N Category:							
N0	51	71.8	11	15.5	40	56.3	
N1	18	25.4	4	5.6	14	19.7	
N2	2	2.8	0	0.0	2	2.8	0.758
M Category:							
M0	49	69.0	10	14.1	39	54.9	
M1	22	31.0	5	7.0	17	23.9	1.000
SDC1 Epithelial Reactivity:							
Absent	57	80.3	11	15.5	46	64.8	
Present	13	18.3	4	5.6	9	12.7	
ND	1	1.4	0	0.0	1	1.4	0.593

5.3.5 Characterisation of Prostate Cancer SDC1 Positive Stromal (PCSP) cells in the Bath Cohort

To try and identify the PCSP cell population, we used double-IF labelling with SDC1 and a range of markers commonly used to distinguish key cell types found in prostate tumour samples. Six tumour samples from the Bath Cohort which had confirmed PCSP cell burden were stained with markers for immune, epithelial and stromal cell lineages.

5.3.5.1 PCSP cells do not express immune cell markers, but are found in proximity to inflammation

Given that SDC1 is expressed in normal and neoplastic plasma cells (Chilosi *et al.*, 1999; Maiga *et al.*, 2014), I began with the hypothesis that PCSP cells were a plasma cell population infiltrating prostate tumours. I began by assessing the expression of lymphoid lineage markers in PCSP cells. Fluorescent co-stains with SDC1 and a panel of antibodies raised against immune cell lineage markers: CD45, which is expressed in the majority of haematolymphoid lineages with the exception of erythrocytes and some macrophages, and known to be expressed in SDC1+ bone marrow plasma cells (Maiga *et al.*, 2014); CD19, which is expressed on immature and memory B-cells (Mavropoulos *et al.*, 2016); and CD27, which is expressed in plasma cells, memory B-cells (Mavropoulos *et al.*, 2016) and T-cells (Hendriks *et al.*, 2000). This showed that the majority of PCSP cells did not co-stain for any of the three immune cell markers (Figure 5.6A and B). Interestingly, despite the lack of colocalisation, PCSP cells were often found in close proximity to CD45 positive immune cells (Figure 5.6A, arrows). This was consistent with IHC data from Chapter 5.3.1 showing that SDC1 positive cells were found in close proximity to cells whose nuclear morphology – round, compact and hyperchromatic - suggested an immune cell phenotype (Figure 5.1B, arrowheads). Due to lack of reactivity with immune cell markers, I proceeded to test the PCSP cell population for colocalisation with markers of other prostatic cell types.


A
CD45
SDC1
DAPI
Merge
CD19
SDC1
DAPI
Merge
CD27
SDC1
DAPI
Merge

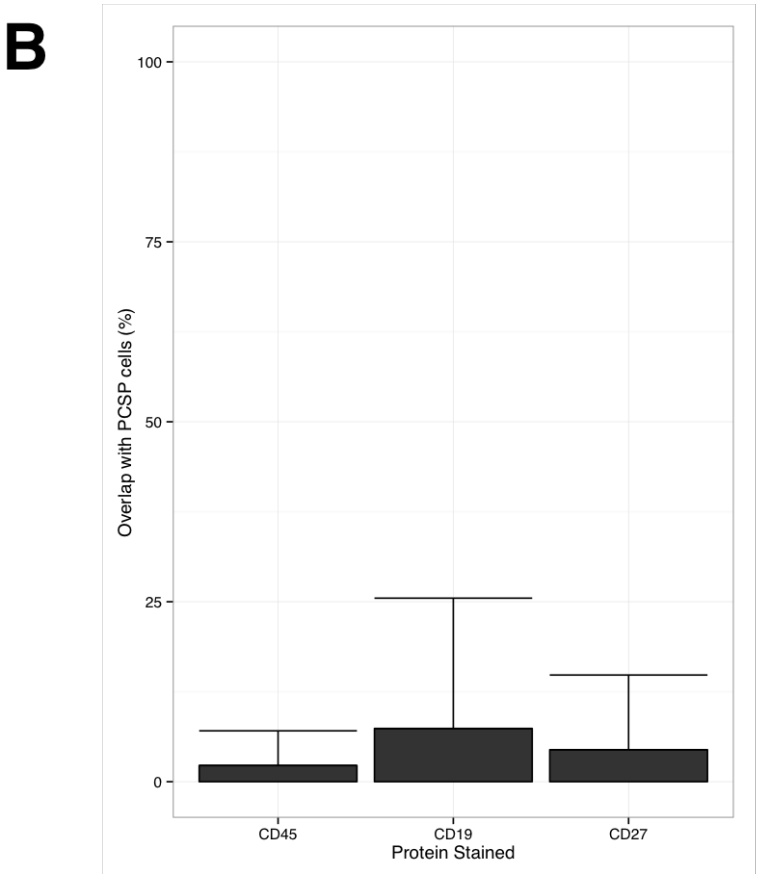

B
100
75
50
25
0
Overlap with PCSP cells (%)
CD45
CD19
CD27
Protein Stained

Figure 5.6: PCSP cells are found in inflammatory areas of prostate tumours, but are not positive for common markers of immune cell lineages.

5µm thick sections of FFPE prostate tumour tissue from the Bath Cohort were stained by IF for SDC1 (green) and immune cell markers (red). Nuclei were counterstained with DAPI (blue). (A) PCSP cells (inset) do not express the hematopoietic lineage marker CD45, the B-cell marker CD19 or the plasma cell marker CD27. (B) Quantification of the percentage overlap of SDC1+ cells with marker stains. Counts of PCSP cells were pooled from a minimum of 3 different patient samples, with a minimum of 200 cells counted in total per marker over at least 5 fields of view. Scale bars – 50µm. Insets: individual PCSP cells displayed at 3x zoom.

5.3.5.2 PCSP cells do not express epithelial cell markers

Another possible origin for the PCSP population was epithelial cells, so co-staining was carried out for SDC1 and three markers of epithelial cell fate: pan-Cytokeratin, which stains multiple human cytokeratins expressed in epithelial tissues (Figure 5.7A); E-Cadherin, which is normally present in the benign prostate epithelium and in most prostate tumours (Rubin *et al.*, 2001); and Prostate-Specific Antigen (PSA), a secretory marker of luminal prostate epithelium, also expressed in prostate tumours. Quantification of the IF images showed none of these markers had significant co-staining with SDC1 in PCSP cells (Figure 5.7B). There was also no apparent overlap in the IHC staining of the stromal cells in consecutive tissue sections for SDC1 and the basal cell markers p63 and Cytokeratin-5 (Figure 5.7C). Thus, PCSP cells did not have staining patterns consistent with them being a recognized prostate epithelial cell population, whether basal or luminal in origin.

5.3.5.3 PCSP cells do not express stromal cell markers

To examine if the PCSP cells were stromal in origin, or had undergone an epithelial-mesenchymal transition and lost expression of epithelial markers, a panel of stains for mesenchymal and stromal cell types were examined. The mesenchymal cell markers Vimentin and N-Cadherin did not show significant co-staining with PCSP cells (Figure 5.8A and B). Finally, the nerve marker S100 and the endothelial marker CD31 were examined, neither of which showed significant overlap (Figure 5.8A and B). Therefore, the marker analysis carried out did not allow the allocation of PCSP cells to an established cell population, and they remain an undefined cell type.

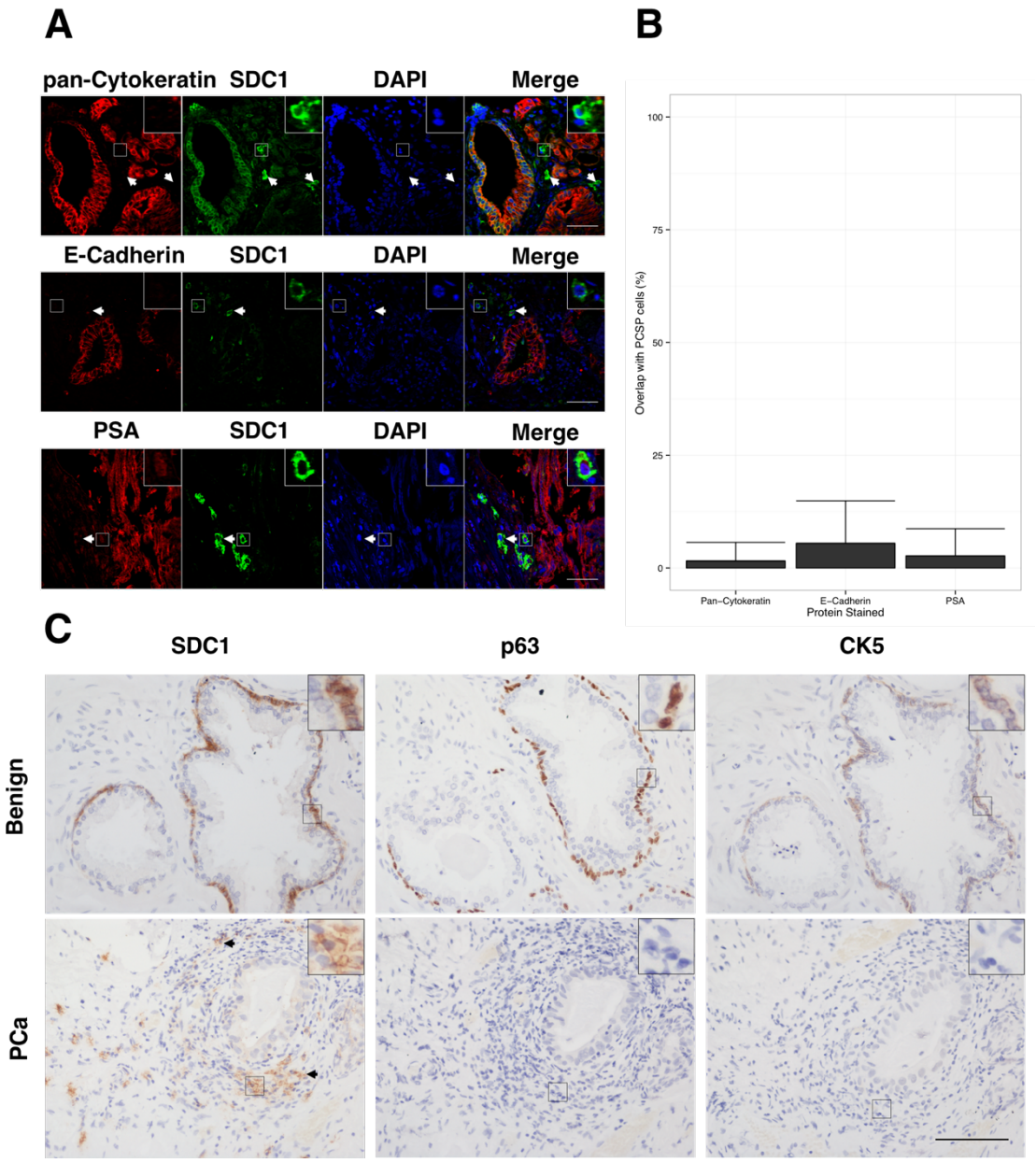
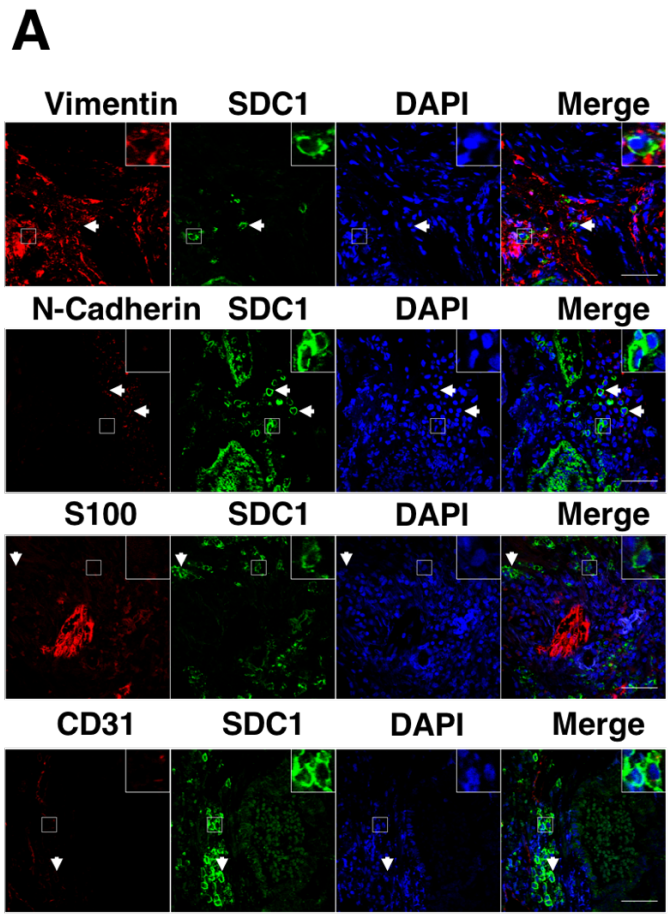

A
pan-Cytokeratin
SDC1
DAPI
Merge
E-Cadherin
SDC1
DAPI
Merge
PSA
SDC1
DAPI
Merge
B
Overlap with PCSP cells (%)
100
75
50
25
0
Pan−Cytokeratin
E−Cadherin
PSA
Protein Stained
C
SDC1
p63
CK5
Benign
PCa

Figure 5.7: PCSP cells do not express markers of epithelial and secretory prostate cells.

5μm thick sections of FFPE prostate tumour tissue from the Bath Cohort were stained by IF for SDC1 (green) and epithelial markers (red). Nuclei were counterstained with DAPI (blue). (A) PCSP cells (inset) do not express the epithelial cell markers Pan-Cytokeratin or E-Cadherin. Similarly, the secretory epithelial marker PSA was also not expressed in PCSP cells. (B) Quantification of the percentage overlap of PCSP cells with marker stains. Counts of PCSP cells were pooled from a minimum of 3 different patient samples. A minimum of 200 cells were counted in total per marker over at least 5 fields of view, with the exception of E-Cadherin where 176 cells were counted. (C) Adjacent sections of benign prostate and prostate tumours were stained for SDC1 and the basal cell markers p63 and CK5, showing no colocalisation in PCSP cells of tumour cases. Insets: individual PCSP cells displayed at 3x zoom. Scale bars – 50μm.


A
Vimentin
SDC1
DAPI
Merge
N-Cadherin
SDC1
DAPI
Merge
S100
SDC1
DAPI
Merge
CD31
SDC1
DAPI
Merge

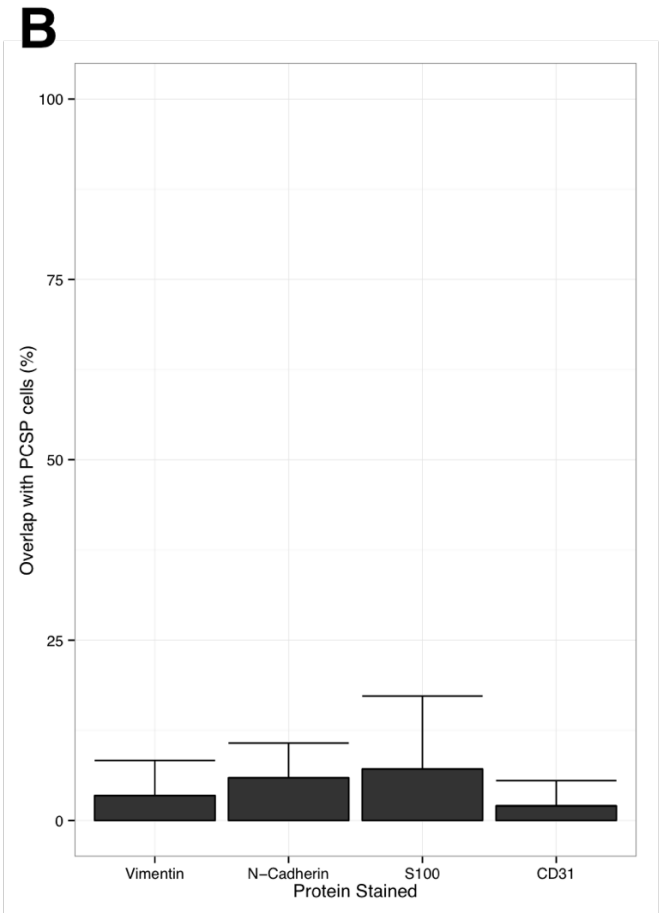

B
100
75
50
25
0
Overlap with PCSP cells (%)
Vimentin
N–Cadherin
S100
CD31
Protein Stained

Figure 5.8: PCSP cells do not express common markers consistent with a mesenchymal/stromal, nerve or endothelial cell identity.

5µm thick sections of FFPE prostate tumour tissue were stained by IF for SDC1 (green) and mesenchymal/stromal cell type markers (red), and nuclei were counterstained with DAPI (blue). (A) mesenchymal cell type markers Vimentin and N-Cadherin, nerve marker S100 and endothelial cell marker CD31. (B) Quantification of the percentage overlap of PCSP cells with marker stains. Counts of PCSP cells were pooled from a minimum of 3 different patient samples. A minimum of 200 cells per marker were counted in total per marker over at least 5 fields of view, with the exception of Vimentin, N-Cadherin and S100, where 69, 197 and 29 cells were counted respectively. Insets: individual PCSP stromal cells displayed at 3x zoom. Scale bars – 50µm.

5.3.5.4 PCSP cells are elongated and have prominent lamellipodia-like structures

In order to characterize this unidentified cell type further, the cell shape of PCSP cells was analysed using anti-SDC1 IHC stained tissue microarrays from two independent cohorts of prostate cancer patients. The circularity and aspect ratios of the PCSP cells was quantified and on average PCSP cells were found to be elongated rather than rounded in shape (Table 5.4). When morphological data from individual cells of both TMA cohorts were pooled, a tail of elongated cells with high aspect ratio and low circularity became apparent (Figure 5.9A). To further characterize this subpopulation of elongated cells, confocal laser scanning microscopy was used to carry out high resolution three-dimensional imaging of immunofluorescence-stained PCSP cells. They were found to have striking lamellipodia-like protrusions (Figure 5.9B, arrows) and polarized cell bodies and nuclei (Figure 5.9B, arrowheads). The elongated shape and prominent protrusions provide further evidence that PCSP cells may be a migratory cell type. This evidence is further compounded by the finding of PCSP cells in adjacent normal prostate tissue in Chapter 5.3.4.

Table 5.4: Summary of shape descriptors and statistics from PCSP cells in both tissue microarray cohorts.

	PR803b		***PR1921***	
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**
Area	72.14	41.31	66.42	35.09
Circularity	0.70	0.13	0.69	0.12
Aspect Ratio	1.83	0.68	1.88	0.62
Roundness	0.60	0.17	0.58	0.16
Solidity	0.88	0.07	0.88	0.06

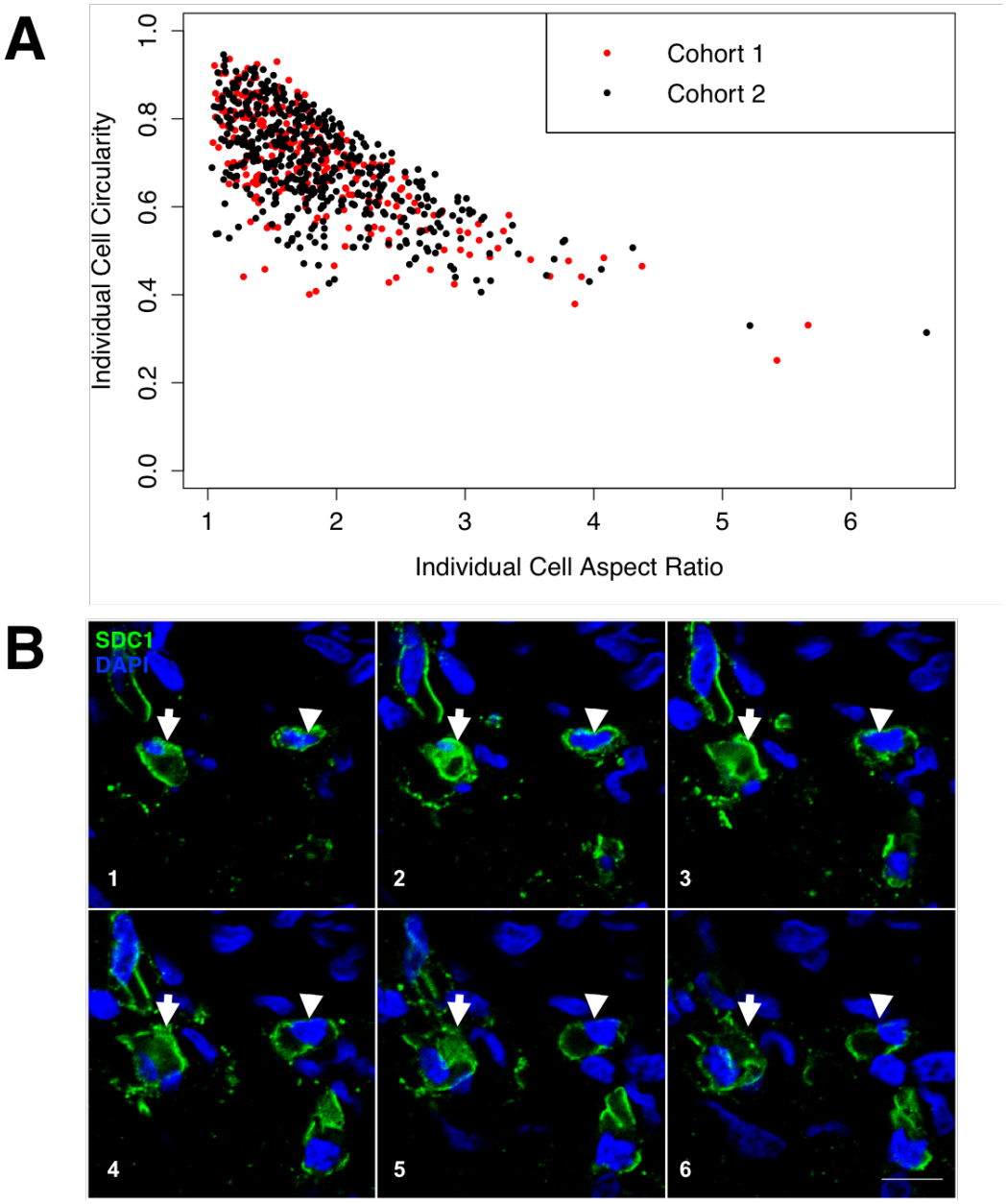

A
Cohort 1
Cohort 2
Individual Cell Circularity
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
Individual Cell Aspect Ratio
B
SDC1
DAPI
1
2
3
4
5
6

Figure 5.9: PCSP cells are elongated and have lamellipodia-like protrusions.

(A) Pooled single cell measurement data of PCSP cells from both tissue microarray cohorts, showing a continuum between a more rounded cell morphology (circularity and aspect ratio close to 1) and a tail of cells assuming a more elongated cell shape with lower circularity and higher aspect ratio was observed. (B) Sequential slices are shown through a *z*-stack of IF-stained PCSP cells in a 5µm thick prostate tumour tissue section, showing elongated cell morphology, polarized nuclei and lamellipodia-like protrusions. Optical sections are 1µm apart. Blue: DAPI; Green: SDC1. Scale bar: 10µm.

5.4 Discussion

By examining SDC1 expression in normal prostate and prostate carcinoma specimens, I confirmed that SDC1 expression is lost in prostate carcinoma. Moreover, SDC1 expression identifies a previously unreported cell population in the stroma of prostate adenocarcinoma with two different anti-SDC1 antibodies in three different patient cohorts and in adjacent normal tissue. These PCSP cells do not express lineage markers of common epithelial, stromal or immune cell lineages, including the plasma cell population which is marked by SDC1 expression (Maiga *et al.*, 2014; Chilosi *et al.*, 1999). PCSP cells display a continuum of rounded and elongated morphologies consistent with a migrating cell type, and appear in high burdens in a subset of primary Gleason grade 5 tumours. Gleason grade 5 morphology is a poor prognostic indicator, making this cell population an interesting but currently unidentified cell type.

5.4.1 Epithelial SDC1 expression is lost in prostate carcinoma but SDC1-expressing stromal cells are present

Previous research on SDC1 expression has identified that the proteoglycan is expressed in basal cells of the benign prostate epithelium and epithelial expression is reduced overall in malignant tissue (Chen *et al.*, 2004a; Shimada *et al.*, 2009; Shariat *et al.*, 2008; Fujii *et al.*, 2015), though some tumours continue to express SDC1 in their epithelial compartments in a cytoplasmic or membranous localization (Chen *et al.*, 2004a; Shariat *et al.*, 2008; Fujii *et al.*, 2015). However, this is controversial, with some studies reporting that there is no difference in epithelial SDC1 expression between benign and tumour tissue (Brimo *et al.*, 2010). This study agrees with the former

findings, with the presence of basal cell staining of SDC1 in benign tissue providing a good positive control for antibody specificity.

In addition to these previously demonstrated patterns, SDC1-expressing PCSP cells were found in the stroma of prostate carcinoma specimens. This is the first time that PCSP cells have been identified in prostate tumours, despite numerous attempts to characterise SDC1 expression patterns in prostate tumours. An *in situ* hybridisation study of multiple epithelial tissues found epithelial but not stromal mRNA localisation in both normal prostate and prostate carcinoma (Mennerich *et al.*, 2004). It is possible that the focal and scattered nature of these cells, and their presence in only a third of surveyed tumour cores by TMA, has caused them to be missed in previous studies. The experimental evidence for the existence of an SDC1-expressing stromal cell population in this work is strong: PCSP cells were found in three independent patient cohorts and with two different anti-SDC1 antibodies, with staining specificity confirmed by double IF on benign and malignant prostate. This makes antibody-specific or fixation-induced staining artifacts an unlikely explanation for the staining pattern.

There may be precedent for stromal expression of SDC1: T47D breast carcinoma cells co-cultured with mammary fibroblasts induce SDC1 expression and ectodomain shedding in the mammary fibroblast compartment, which is required by the carcinoma cells for a proliferative response to FGF2 (Su *et al.*, 2007). This suggests that autocrine or paracrine signalling between carcinoma cells and their stroma can result in SDC1 expression and shedding in the stromal compartment, at least in the breast. Evidence for an epithelial-stromal interaction that promotes SDC1 stromal expression is supported by the finding that breast, lung and colon tissues switch from an epithelial to a stromal pattern of SDC1 expression in carcinoma (Mennerich *et al.*, 2004). However, this has yet to be demonstrated in prostate carcinoma. To more thoroughly demonstrate the existence of PCSP cells in prostate tumours, it would be necessary to isolate them from fresh tissue and characterise them in a tissue culture system *in vitro*, or produce patient-derived xenografts from confirmed PCSP-bearing tumours to study the phenomenon further in mice.

5.4.2 PCSP cell burden may indicate a non-recurrent tumour in the Bath cohort

When the epithelial and stromal staining patterns of SDC1 staining in prostate tumours were quantified and associated with clinical data, I found that SDC1 epithelial staining

has no effect on recurrence status or Gleason grade. This is consistent with several previous reports on SDC1 expression by IHC which suggest it is not associated with recurrence (Purushothaman *et al.*, 2010; Brimo *et al.*, 2010) or with Gleason grade (Brimo *et al.*, 2010). In contrast, PCSP cell burden showed a trend towards being positively associated with non-recurrent tumours, suggesting that the presence of these cells may be associated with favourable outcome. Other infiltrating cells in tumours, such as mast cells, have been previously associated with favourable outcome in prostate cancer patients (Fleischmann *et al.*, 2009). However, given the small size of the Bath Cohort, further examination of a larger patient cohort would be required to verify this finding.

5.4.3 Primary Gleason grade 5 tumours are enriched in PCSP cell burden

Interestingly, we observed that there is a trend towards increased burden in the primary Gleason grade 5 tumours, suggesting that it might also be related to outcome given that assignment of a high Gleason grade is a predictor of metastasis and thus poor prognosis (Antonarakis *et al.*, 2012). While cohort 2 showed a significant association between primary Gleason grade and the presence of a PCSP cell burden, cohort 1 did not reach significance. Meanwhile, the Bath Cohort did not contain any patients with primary Gleason grade 5 tumours, and yet all tumours showed some evidence of PCSP cell burden, even if small. These discrepancies could be due to differences in sampling strategy between the cohorts. In the case of tissue microarrays, the two studies had different core numbers and diameters, with cohort 1 having duplicate 1mm cores and cohort 2 having single 1.5mm cores. In the case of the Bath Cohort, core needle biopsies and prostatectomy specimens acquire far more tissue by area than could be represented in a single TMA core. This could lead to differences in successful detection of SDC1+ cells in tumours given their scattered nature. The utility of PCSP cells as a biomarker could be clarified by conducting a larger study focused on detecting PCSP cells in core needle biopsies, where many small cores can be taken for analysis to investigate the distribution and detection rate of cells and its impact on biomarker efficacy.

5.4.4 PCSP cells are a cryptic cell type in prostate tumours

By using multiplex immunofluorescence, we looked for co-staining of classical epithelial, mesenchymal, stromal and immune markers with the PCSP cell population.

We found no significant overlap between any classical lineage marker tested and PCSP cells, indicating a currently unidentified cell type. Several studies have specifically reported the absence of stromal SDC1 staining in prostate tumours (Chen *et al.*, 2004a; Szarvas *et al.*, 2016b; Shariat *et al.*, 2008). We are also unable to detect a widespread stromal staining pattern in either benign prostate or prostate tumour tissue. PCSP cells do not show common markers of stromal cell types, and it is possible that the focal nature of this cell type – and their presence in a subset of tumours - has made this phenotype more difficult to detect. However, two separate antibodies recognizing SDC1 have reliably detected this cell population - in both immunohistochemistry and immunofluorescence and using three different patient cohorts.

I considered that the PCSP cell population could be immune cells, given their observed proximity to inflammation. SDC1 is a marker of plasma cell differentiation, being expressed at the later stages of the transition from plasmablast into a mature plasma cell (De Vos *et al.*, 2006). Using three markers that would encompass a range of immune cell lineages including B- and T-cell repertoires, we were unable to identify an immune cell marker profile in PCSP cells. Given the enormous diversity in the immune cell repertoire, it is still possible that PCSP cells represent an immune cell population. Chronic inflammation is commonplace in the prostate (Nickel *et al.*, 2008) and clusters of lymphocytes have been documented to surround prostate tumour tissue previously, such as the FOXP3+/PD-1+/PD-L1+ population of regulatory T-cells reported by Ebelt *et al.* (2009), and other studies have observed the infiltration of cytotoxic and regulatory T-cells following androgen deprivation therapy (Mercader *et al.*, 2001; Sorrentino *et al.*, 2011). Additionally, mast cells have been shown to infiltrate into the stroma of prostate tumours (Fleischmann *et al.*, 2009). However, the fact that immune cells tend to express both SDC1 and the pan-leukocyte marker CD45 (Maiga *et al.*, 2014), whilst PCSP cells are CD45-, would suggest that this is an atypical immune cell or non-immune cell population. Tumour-associated macrophages - a component commonly associated with inflammation in tumours – also express CD45 being a leukocyte population, and that makes it unlikely that PCSP cells represent a subset of that population either.

I next examined whether PCSP cells could be epithelial cells (basal, luminal or otherwise). Co-staining with epithelial markers again failed to mark a significant portion of the PCSP cell population. One study has shown a positive correlation between the expression of epithelial SDC1 and E-cadherin and β-catenin, with a trend towards decreased expression as Gleason grade increases (Contreras *et al.*, 2010). There

remains the possibility that PCSP cells could have undergone dedifferentiation, as they do not possess an epithelial identity consistent with tumour cells. Epithelial-Mesenchymal Transition (EMT) is an unlikely explanation for this observation, as PCSP cells also lacked mesenchymal markers such as Vimentin and N-Cadherin. Another possibility for the lack of differentiation markers is a stem-like status, either derived from a resident prostate cell population or from a mobilized stem cell niche such as Mesenchymal Stem Cells (Cheng *et al.*, 2016) or cancer stem cells. There is evidence that SDC1 signalling is involved in maintaining the tumour-initiating cell (TIC)/cancer stem cell (CSC) population in the prostate carcinoma cell line PC-3, where its expression marks a population of CD133+/CD44+ stem-like holoclones (Shimada *et al.*, 2013). Additionally, putative CSCs from the 22Rv1 and DU145 prostate cancer cell lines lack expression of basal or luminal differentiation markers, suggesting that a dedifferentiated CSC-like cell population might not express differentiated cell type markers and thus evade our method of cell type characterisation. Testing this hypothesis in PCSP cells would require extensive co-staining with stem cell markers and/or functional assays with PCSP cells isolated *in vitro*.

Another possibility is that SDC1 is being shed from other cell types in the prostate and bound or taken up by PCSP cells, which cannot yet be excluded because the anti-SDC1 antibody clone H-174 also recognises the soluble ectodomain, and the binding epitope for the B-A38 antibody is unknown. Additionally, while SDC1 expression is lost in the epithelial compartment, some residual heterogeneous SDC1 staining is often observed which could potentially correspond to the shed form of the protein being produced in tumour cells. Shed SDC1 from tumour cells has been shown to enter stromal cells in a non-cell autonomous manner previously (Stewart *et al.*, 2015). Future attempts to distinguish whether the recognised protein on the surface of PCSP cells is shed or full-length would require staining with antibodies specific to the soluble ectodomain and the core protein.

5.4.5 The PCSP cell population is also present in adjacent normal tissue and may contain migratory cells

Our data suggest that, while PCSP cells are predominantly present in a subset of prostate tumours, histologically normal tissue adjacent to a carcinoma can also harbour PCSP cells. This would be consistent with the hypothesis of a migratory cell population either recruited from other tissues, or migrating out from the prostate tumour. Detection

of PCSP cells in adjacent normal tissue could expand the diagnostic and prognostic utility of the biomarker in a biopsy setting, where many small cores are taken and histologically normal tissue is a common finding. It is possible that, if the PCSP cell type is only present in benign tissue that is adjacent to a carcinoma, it may be possible to both confirm diagnosis and glean prognostic information from even a normal tissue core. Additionally, a recent study has demonstrated a prognostically relevant field effect at the gene expression level that can be detected in adjacent normal tissue to prostate tumours (Magi-Galluzzi *et al.*, 2016), suggesting that using such tissue for prognostication could be feasible if suitable biomarkers were identified. A larger study with a larger proportion of matched normal adjacent tissue would be required to further assess the diagnostic and prognostic importance of PCSP cell presence in both benign and malignant tissue settings.

The hypothesis of PCSP cells being migratory is further supported by morphological data, suggesting a continuum of rounded and elongated cells. SDC1 can be shed and it is possible that the protein being detected is the shed version acting in a non-cell-autonomous manner, rather than expressed by PCSP cells. To characterize the behaviour of these cells in more detail, migration and chemotaxis assays would need to be performed on live cells from fresh patient tissue. Alternatively, more information about PCSP cell morphology could be obtained *in situ* by co-staining for components of the actin cytoskeleton and focal adhesions in prostate tumours.

5.4.6 Conclusions

Although the prognostic importance of tissue (Zellweger *et al.*, 2003; Chen *et al.*, 2004a; Shariat *et al.*, 2008; Shimada *et al.*, 2009; Brimo *et al.*, 2010; Contreras *et al.*, 2010; Ledezma *et al.*, 2011; Poblete *et al.*, 2014) and serum SDC1 (Szarvas *et al.*, 2016b) has been assessed by multiple studies and remains controversial, the prognostic significance of PCSP cells had not been elucidated. The current study is the first to identify stromally-located SDC1+ cells, an unknown cell type, found in prostate carcinoma and its adjacent normal tissue, and their heightened burden in Gleason pattern 5 tumours. This provides the foundation for a detailed investigation into the identity - and diagnostic and prognostic utility - of PCSP cells in prostate carcinoma.

6 Final Discussion

Prostate cancer is a challenging disease to diagnose and prognosticate, as many patients present with disease that may not be a significant risk to health, whereas a subset of approximately 20-25% of patients will experience recurrence following primary therapy. Gleason grade and tumour stage are currently the best prognostic biomarkers available, and biomarkers that better predict patient outcome are urgently needed. The current project aimed to identify new biomarkers to predict disease recurrence in primary tumour samples from patients with prostate cancer, in attempt to risk-stratify patients and produce more tailored therapies.

6.1 Main conclusions

This thesis presents the following key findings:

- Potential biomarkers related to stem cells and cancer stem cells could be identified using a literature searching approach, with SDC1, BMI1, MUC1-C, ALDH7A1, Nestin and ZSCAN4 being selected for evaluation by IHC. Out of these, heparan sulphate proteoglycan SDC1 and pluripotency factor ZSCAN4 show novel staining patterns in prostate tumours that warrant further investigation.
- A bioinformatic approach identified secreted matrix protein RS1 and SLC31A1 as potential biomarkers. As well as being overexpressed at the mRNA level in relapsing tumours, the RS1 gene is amplified in up to 32% of metastatic and castration-resistant prostate tumours. Bath cohort evidence suggests that RS1 protein expression may be elevated in recurrent primary prostate tumours at the protein level, and negatively correlated with AR protein expression.
- The extremes of expression analysis method was developed to study the potential functional consequences arising from elevated expression of genes of interest at the transcriptional level. The method confirms already known functions of SDC1 in prostate cancer transcriptome datasets as well as identifying hypothetical new functions for ZSCAN4 in metabolism.
- SDC1 is expressed in a novel cell type, referred to as Prostate Cancer SDC1-Positive Stromal (PCSP) cells, in the stroma of approximately a third of primary tumours and in adjacent normal tissue. These cells do not express markers of immune cell, epithelial or mesenchymal differentiation, making them an unidentified cell type. Additionally, PCSP cells show evidence of lamellipodia-like structures and an elongated cell shape suggestive of a migratory cell type, although this requires further investigation.

6.2 Biomarker identification methods were effective but could be improved

This project utilised both a manually-curated literature searching approach and a bioinformatic approach to identify biomarkers related to recurrence and cancer stem cells (Chapter 3). Interestingly, although every potential biomarker identified by literature searching had been reported to be expressed in putative CSCs with the exception of ZSCAN4, SDC1 and ZSCAN4 were the only biomarkers to show restricted patterns of protein expression to a subset of epithelial or stromal cells (Chapter 4). The rest of the markers showed widespread staining in glandular epithelium, suggesting that they alone do not identify a stem-like fraction within prostate tumours. Stem-like cell populations would be expected to constitute only a small fraction of epithelial tissues, and the CSC population of prostate cancer cell lines and primary tumours is similarly restricted to a subpopulation of cells characterised by the expression of multiple markers or properties (see Chapter 1.3). For example, when the pluripotency marker ZSCAN4 was originally identified, whole mount *in situ* hybridisation found that it had a heterogeneous and restricted expression pattern in mouse embryonic stem cell colonies (Falco *et al.*, 2007; Carter *et al.*, 2008). Therefore, while the literature searching approach has identified several potential biomarkers of interest in prostate cancer, it does not appear to identify CSCs in this tissue.

One disadvantage of the approaches taken in this project, where candidates are filtered and tested by IHC, is that only a small number of proteins can be examined. To identify CSC markers related to prostate cancer, I speculate that an experimental screening approach could be invaluable when combined with the initial candidate biomarker approach. However, antibody staining by IHC often requires optimisation of reaction conditions, which was often difficult to achieve for many of our antibodies in the mouse TMA constructed (Chapter 3), and prostate biopsies are too small and precious to use for optimisation purposes. A high-throughput screen of many proteins by IHC in prostate tissue is therefore unfeasible. In contrast, high-throughput expression screening by *in situ* hybridisation is readily used for heterogeneous mRNA expression pattern detection in embryonic stem cells (Carter *et al.*, 2008) and even in model organisms such as Medaka embryos (Quiring *et al.*, 2004). *In situ* hybridisation screening could be conducted on explanted prostate tumour tissue or prostate cancer cell lines in a high-throughput setting, using a panel of probes recognising previously identified stem cell markers, candidate genes such as those identified by literature searching, or even probes

generated from a human embryonic stem cell cDNA library. Heterogeneous or restricted expression patterns could then be further characterised *in vitro* and in human prostate tumour tissue to identify candidate markers for putative CSCs. Such an approach would provide additional experimental validation before potential biomarkers are investigated in depth.

It is important to note that, despite only 7 candidates being tested by IHC in total, both approaches yielded potential biomarkers of interest in prostate cancer, suggesting that the literature searching and bioinformatics methods are effective in identifying potential candidates. Examples include: a heightened burden of SDC1-expressing PCSP cells being related to high-grade tumours, ZSCAN4 being heterogeneously expressed in a subpopulation of prostate cancer patients, and RS1 being associated with recurrence. The main limitation of this approach was in sample size, as the Bath Cohort was not large enough to fully test the relationships between potential biomarkers and the clinical data. Future work for characterisation of these markers should therefore include larger patient cohorts, likely using tissue microarray technology, and should also include important clinical data such as recurrence-free survival and overall survival.

While many markers of putative CSCs in the human prostate are related to tumorigenicity and metastasis in mouse xenograft models, it remains to be seen if a population of CSCs exist in the human prostate tumour *in situ* (see Chapter 1.3). This project found biomarkers with relationships to clinical variables but did not find any biomarker expression patterns that were consistent with a CSC niche in prostate tumours. Therefore, the question of whether CSCs are related to recurrence of prostate carcinoma still remains to be addressed. Regardless of the existence of prostate CSCs, the phenotypes associated with putative prostate CSC populations such as CD44-expressing cells, including an increased propensity for proliferation and metastasis (Patrawala *et al.*, 2006), and the ability of these markers to regulate these phenotypes (Liu *et al.*, 2011a), both suggest a relationship with prostate cancer cell behaviour that may be prognostically relevant.

6.3 Integration of data from the DNA, RNA and protein levels is essential for characterising potential biomarkers in prostate tumours

The landscape of prostate cancer is marked by many different mutational events, from its potential beginnings in PIA or PIN through to carcinoma and castration-resistant disease (Chapter 1.2.3). These mutations, which can be due to point mutations, rearrangements, losses and gains of individual loci or even whole chromosomes and chromosomal arms, are thought to be key for driving disease progression and may even be predictive of recurrence (Hieronymus *et al.*, 2014). There have been several attempts to profile the genomes and transcriptomes of both primary prostate tumours (Tomlins *et al.*, 2007; Abeshouse *et al.*, 2015; Ellwood-Yen *et al.*, 2003; Grasso *et al.*, 2012; Taylor *et al.*, 2010) and castration-resistant disease (Tomlins *et al.*, 2007; Grasso *et al.*, 2012; Robinson *et al.*, 2015) in order to identify key drivers and signalling pathways involved in prostate cancer progression and therapeutic resistance. Harnessing this wealth of data at the DNA and RNA level has merit for both identifying potentially novel biomarkers and for filtering and further characterising biomarkers identified by other means.

A good example of this is RS1, which was identified as being upregulated at the mRNA level in recurrent prostate tumours in a cDNA microarray dataset (Chapter 3). This trend was also apparent at the protein level by IHC as well as a negative correlation between RS1 protein levels and AR protein levels, although a larger patient cohort is required to verify these findings (Chapter 4). When combined with data from a meta-analysis of large-scale prostate cancer genomics studies, which found that RS1 is amplified in castration-resistant prostate carcinomas, the integrative approach creates a stronger case for continued study of RS1 in prostate cancer as a potential prognostic biomarker.

Another example that supports the use of an integrative approach comes from the study of SDC1, where the extremes of expression method adds additional weight to the hypotheses generated as a result of the data generated (Chapter 5). I found a novel cell type expressing SDC1, PCSP cells, in prostate tumour tissue and its adjacent normal tissue, but these cells still remain unidentified despite attempts to identify their differentiation markers. Morphological examination of PCSP cells suggested that they may be a migratory cell type, but this has yet to be confirmed. Extremes of expression analysis on SDC1 in two prostate cancer cDNA microarray datasets (Chapter 3) provided additional information on potential roles of SDC1 expression in prostate tumours: GO terms associated with cell motility and cell adhesion and the acute

immune response were enriched in patients with high expression of SDC1 in their tumours. Given that SDC1 expression is often lost in prostate carcinoma, and that these transcriptomics datasets measure the transcriptome from all components of the tumour tissue, it is tempting to speculate that this expression pattern and its GO terms are reflecting PCSP cells. Moreover, the proximity of PCSP cells to inflammation and their elongated cell shapes with observed lamellipodia-like structures *in vivo* (Chapter 5) are further evidence to back up the observations from the extremes of expression analysis *in silico*. These combined findings also inform the direction of future experiments, which should aim to identify the cell type marked by PCSP cells, and additionally characterise their biological properties – including cell migration - *in vitro* and *in vivo*.

6.4 Final conclusion

In summary, the current project attempted to identify new prognostic biomarkers for prostate cancer. The work presented in this thesis suggests that, while a CSC-related candidate biomarker approach is insufficient to identify CSCs in prostate tumours, both candidate biomarker and bioinformatics approaches are suitable for identifying potential biomarkers for prostate cancer. 7 potential biomarkers were identified using literature searching and bioinformatics techniques, further characterised using bioinformatics analysis of microarray and genomics data, and subjected to evaluation by IHC. SDC1 marked an unidentified cell type found in the stroma of Gleason pattern 5 tumours; RS1 was potentially upregulated in recurrent prostate tumours and amplified in castration-resistant disease; and ZSCAN4 had a heterogeneous expression pattern that was present in a subset of prostate tumours with as yet undefined clinical and biological characteristics. Therefore, at least 3 of 7 potential biomarkers warrant more extensive investigation in prostate tumours.

7 REFERENCES

Abate-Shen, C. & Shen, M.M., 2000. Molecular genetics of prostate cancer. *Genes & Development,* 14(19), pp. 2410-2434.

Abate-Shen, C., Shen, M.M. & Gelmann, E., 2008. Integrating differentiation and cancer: The Nkx3.1 homeobox gene in prostate organogenesis and carcinogenesis. *Differentiation,* 76(6).

Abdulkadir, S.A., Magee, J.A., Peters, T.J., Kaleem, Z., Naughton, C.K., Humphrey, P.A. & Milbrandt, J., 2002. Conditional loss of Nkx3.1 in adult mice induces prostatic intraepithelial neoplasia. *Molecular and Cellular Biology,* 22(5), pp. 1495-1503.

Abdullah, L.N. & Chow, E.K.-H., 2013. Mechanisms of chemoresistance in cancer stem cells. *Clinical and translational medicine,* 2(1), pp. 3-3.

Abeshouse, A., Ahn, J., Akbani, R., Ally, A., Amin, S., Andry, C.D., Annala, M., Aprikian, A., Armenia, J., Arora, A., Auman, J.T., Balasundaram, M., Balu, S., Barbieri, C.E., Bauer, T., Benz, C.C., Bergeron, A., Beroukhim, R., Berrios, M., Bivol, A., Bodenheimer, T., Boice, L., Bootwalla, M.S., dos Reis, R.B., Boutros, P.C., Bowen, J., Bowlby, R., Boyd, J., Bradley, R.K., Breggia, A., Brimo, F., Bristow, C.A., Brooks, D., Broom, B.M., Bryce, A.H., Bubley, G., Burks, E., Butterfield, Y.S.N., Button, M., Canes, D., Carlotti, C.G., Carlsen, R., Carmel, M., Carroll, P.R., Carter, S.L., Cartun, R., Carver, B.S., Chan, J.M., Chang, M.T., Chen, Y., Cherniack, A.D., Chevalier, S., Chin, L., Cho, J., Chu, A., Chuah, E., Chudamani, S., Cibulskis, K., Ciriello, G., Clarke, A., Cooperberg, M.R., Corcoran, N.M., Costello, A.J., Cowan, J., Crain, D., Curley, E., David, K., Demchok, J.A., Demichelis, F., Dhalla, N., Dhir, R., Doueik, A., Drake, B., Dvinge, H., Dyakova, N., Felau, I., Ferguson, M.L., Frazer, S., Freedland, S., Fu, Y., Gabriel, S.B., Gao, J., Gardner, J., Gastier-Foster, J.M., Gehlenborg, N., Gerken, M., Gerstein, M.B., Getz, G., Godwin, A.K., Gopalan, A., Graefen, M., Graim, K., Gribbin, T., Guin, R., Gupta, M., Hadjipanayis, A., Haider, S., Hamel, L., Hayes, D.N., Heiman, D.I*., et al.*, 2015. The Molecular Taxonomy of Primary Prostate Cancer. *Cell,* 163(4), pp. 1011-1025.

Ahearn, T.U., Pettersson, A., Ebot, E.M., Gerke, T., Graff, R.E., Morais, C.L., Hicks, J.L., Wilson, K.M., Rider, J.R., Sesso, H.D., Fiorentino, M., Flavin, R., Finn, S., Giovannucci, E.L., Loda, M., Stampfer, M.J., De Marzo, A.M., Mucci, L.A. & Lotan, T.L., 2016. A Prospective Investigation of PTEN Loss and ERG Expression in Lethal Prostate Cancer. *Jnci-Journal of the National Cancer Institute,* 108(2).

Ahlgren, G., Pedersen, K., Lundberg, S., Aus, G., Hugosson, J. & Abrahamsson, P.A., 2000. Regressive changes and neuroendocrine differentiation in prostate cancer after neoadjuvant hormonal treatment. *Prostate,* 42(4), pp. 274-279.

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. & Clarke, M.F., 2003. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America,* 100(7), pp. 3983-3988.

Albino, D., Civenni, G., Dallavalle, C., Roos, M., Jahns, H., Curti, L., Rossi, S., Pinton, S., D'Ambrosio, G., Sessa, F., Hall, J., Catapano, C.V. & Carbone, G.M., 2016. Activation of the Lin28/let-7 Axis by Loss of ESE3/EHF Promotes a Tumorigenic and Stem-like Phenotype in Prostate Cancer. *Cancer Research,* 76(12), pp. 3629-3643.

Albino, D., Longoni, N., Curti, L., Mello-Grand, M., Pinton, S., Civenni, G., Thalmann, G., D'Ambrosio, G., Sarti, M., Sessa, F., Chiorino, G., Catapano, C.V. & Carbone, G.M., 2012. ESE3/EHF Controls Epithelial Cell Differentiation and Its Loss Leads to Prostate Tumors with Mesenchymal and Stem-like Features. *Cancer Research,* 72(11), pp. 2889-2900.

Alexander, C.M., Reichsman, F., Hinkes, M.T., Lincecum, J., Becker, K.A., Cumberledge, S. & Bernfield, M., 2000. Syndecan-1 is required for Wnt-1-induced mammary tumorigenesis in mice. *Nature Genetics,* 25(3), pp. 329-332.

Alison, M.R., Guppy, N.J., Lim, S.M.L. & Nicholson, L.J., 2010. Finding cancer stem cells: are aldehyde dehydrogenases fit for purpose? *Journal of Pathology,* 222(4), pp. 335-344.

Alkema, M.J., Jacobs, H., vanLohuizen, M. & Berns, A., 1997. Perturbation of B and T cell development and predisposition to lymphomagenesis in E mu Bmi1 transgenic mice require the Bmi1 RING finger. *Oncogene,* 15(8), pp. 899-910.

Allred, D.C., Harvey, J.M., Berardo, M. & Clark, G.M., 1998. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Modern Pathology,* 11(2), pp. 155-168.

Allsbrook, W.C., Mangold, K.A., Johnson, M.H., Lane, R.B., Lane, C.G. & Epstein, J.I., 2001. Interobserver reproducibility of Gleason grading of prostatic carcinoma: General pathologists. *Human Pathology,* 32(1), pp. 81-88.

Altemeier, W.A., Schlesinger, S.Y., Buell, C.A., Brauer, R., Rapraeger, A.C., Parks, W.C. & Chen, P., 2012. Transmembrane and Extracellular Domains of Syndecan-1 Have Distinct Functions in Regulating Lung Epithelial Migration and Adhesion. *Journal of Biological Chemistry,* 287(42), pp. 34927-34935.

Amano, T., Hirata, T., Falco, G., Monti, M., Sharova, L.V., Amano, M., Sheer, S., Hoang, H.G., Piao, Y., Stagg, C.A., Yamamizu, K., Akiyama, T. & Ko, M.S.H., 2013. Zscan4 restores the developmental potency of embryonic stem cells. *Nature communications,* 4, pp. 1966-1966.

Ammirante, M., Kuraishy, A.I., Shalapour, S., Strasner, A., Ramirez-Sanchez, C., Zhang, W.Z., Shabaik, A. & Karin, M., 2013. An IKK alpha-E2F1-BMI1 cascade activated by infiltrating B cells controls prostate regeneration and tumor recurrence. *Genes & Development,* 27(13), pp. 1435-1440.

An, J., Wang, C.J., Deng, Y.B., Yu, L. & Huang, H.J., 2014. Destruction of Full-Length Androgen Receptor by Wild-Type SPOP, but Not Prostate-Cancer-Associated Mutants. *Cell Reports,* 6(4), pp. 657-669.

Antonarakis, E.S., Feng, Z.Y., Trock, B.J., Humphreys, E.B., Carducci, M.A., Partin, A.W., Walsh, P.C. & Eisenberger, M.A., 2012. The natural history of metastatic progression in men with prostate-specific antigen recurrence after radical prostatectomy: long-term follow-up. *Bju International,* 109(1), pp. 32-39.

Ashida, S., Nakagawa, H., Katagiri, T., Furihata, M., Iiizumi, M., Anazawa, Y., Tsunoda, T., Takata, R., Kasahara, K., Miki, T., Fujioka, T., Shuin, T. & Nakamura, Y., 2004. Molecular features of the transition from prostatic intraepithelial neoplasia (PIN) to prostate cancer: Genome-wide gene-expression profiles of prostate cancers and PINs. *Cancer Research,* 64(17), pp. 5963-5972.

Augustin, H., Mayrhofer, K., Pummer, K. & Mannweiler, S., 2013. Relationship between prostate cancer gene 3 (PCA3) and characteristics of tumor aggressiveness. *Prostate,* 73(2), pp. 203-210.

Bao, S.D., Wu, Q.L., McLendon, R.E., Hao, Y.L., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D. & Rich, J.N., 2006. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature,* 444(7120), pp. 756-760.

Barbieri, C.E., Baca, S.C., Lawrence, M.S., Demichelis, F., Blattner, M., Theurillat, J.-P., White, T.A., Stojanov, P., Van Allen, E., Stransky, N., Nickerson, E., Chae, S.-S., Boysen, G., Auclair, D., Onofrio, R.C., Park, K., Kitabayashi, N., MacDonald, T.Y., Sheikh, K., Vuong, T., Guiducci, C., Cibulskis, K., Sivachenko, A., Carter, S.L., Saksena, G., Voet, D., Hussain, W.M., Ramos, A.H., Winckler, W., Redman, M.C., Ardlie, K., Tewari, A.K., Mosquera, J.M., Rupp, N., Wild, P.J., Moch, H., Morrissey, C., Nelson, P.S., Kantoff, P.W., Gabriel, S.B., Golub, T.R., Meyerson, M., Lander, E.S., Getz, G., Rubin, M.A. & Garraway, L.A., 2012. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nature Genetics,* 44(6), pp. 685-U107.

Barboro, P., Salvi, S., Rubagotti, A., Boccardo, S., Spina, B., Truini, M., Carmignani, G., Introini, C., Ferrari, N., Boccardo, F. & Balbi, C., 2014. Prostate cancer: Prognostic significance of the association of heterogeneous nuclear ribonucleoprotein K and androgen receptor expression. *International Journal of Oncology,* 44(5), pp. 1589-1598.

Barfeld, S.J., Fazli, L., Persson, M., Marjavaara, L., Urbanucci, A., Kaukoniemi, K.M., Rennie, P.S., Ceder, Y., Chabes, A., Visakorpi, T. & Mills, I.G., 2015. Myc-dependent purine biosynthesis affects nucleolar stress and therapy response in prostate cancer. *Oncotarget,* 6(14), pp. 12587-12602.

Barwick, B.G., Abramovitz, M., Kodani, M., Moreno, C.S., Nam, R., Tang, W., Bouzyk, M., Seth, A. & Leyland-Jones, B., 2010. Prostate cancer genes associated with TMPRSS2-ERG gene fusion and prognostic of biochemical recurrence in multiple cohorts. *British Journal of Cancer,* 102(3), pp. 570-576.

Beauvais, D.M. & Rapraeger, A.C., 2010. Syndecan-1 couples the insulin-like growth factor-1 receptor to inside-out integrin activation. *Journal of Cell Science,* 123(21), pp. 3796-3807.

Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A. & Weinberg, R.A., 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nature Genetics,* 40(5), pp. 499-507.

Benjamini, Y. & Hochberg, Y., 1995. CONTROLLING THE FALSE DISCOVERY RATE - A PRACTICAL AND POWERFUL APPROACH TO MULTIPLE TESTING. *Journal of the Royal Statistical Society Series B-Methodological,* 57(1), pp. 289-300.

Bensalah, K., Montorsi, F. & Shariat, S.F., 2007. Challenges of cancer biomarker profiling. *European Urology,* 52(6), pp. 1601-1609.

Berger, M.F., Lawrence, M.S., Demichelis, F., Drier, Y., Cibulskis, K., Sivachenko, A.Y., Sboner, A., Esgueva, R., Pflueger, D., Sougnez, C., Onofrio, R., Carter, S.L., Park, K., Habegger, L., Ambrogio, L., Fennell, T., Parkin, M., Saksena, G., Voet, D., Ramos, A.H., Pugh, T.J., Wilkinson, J., Fisher, S., Winckler, W., Mahan, S., Ardlie, K., Baldwin, J., Simons, J.W., Kitabayashi, N., MacDonald, T.Y., Kantoff, P.W., Chin, L., Gabriel, S.B., Gerstein, M.B., Golub, T.R., Meyerson, M., Tewari, A., Lander, E.S., Getz, G., Rubin, M.A. & Garraway, L.A., 2011. The genomic complexity of primary human prostate cancer. *Nature,* 470(7333).

Bethel, C.R. & Bieberich, C.J., 2007. Loss of Nkx3.1 expression in the Transgenic adenocarcinoma of mouse prostate model. *Prostate,* 67(16), pp. 1740-1750.

Bethel, C.R., Faith, D., Li, X., Guan, B., Hicks, J.L., Lan, F.S., Jenkins, R.B., Bieberich, C.J. & De Marzo, A.M., 2006. Decreased NKX3.1 protein expression in focal prostatic atrophy, prostatic Intraepithelial neoplasia, and adenocarcinoma: Association with Gleason score and chromosome 8p deletion. *Cancer Research,* 66(22), pp. 10683-10690.

Bianco, F.J., Jr., Wood, D.P., Jr., Cher, M.L., Powell, I.J., Souza, J.W. & Pontes, J.E., 2003. Ten-year survival after radical prostatectomy: specimen Gleason score is the predictor in organ-confined prostate cancer. *Clinical prostate cancer,* 1(4), pp. 242-7.

Bismar, T.A., Yoshimoto, M., Duan, Q.L., Liu, S.H., Sircar, K. & Squire, J.A., 2012. Interactions and relationships of PTEN, ERG, SPINK1 and AR in castration-resistant prostate cancer. *Histopathology,* 60(4), pp. 645-652.

Blackwood, J.K., Williamson, S.C., Greaves, L.C., Wilson, L., Rigas, A.C., Sandher, R., Pickard, R.S., Robson, C.N., Tumbull, D.M., Taylor, R.W. & Heer, R., 2011. In situ lineage tracking of human prostatic epithelial stem cell fate reveals a common clonal origin for basal and luminal cells. *Journal of Pathology,* 225(2), pp. 181-188.

Blum, R., Gupta, R., Burger, P.E., Ontiveros, C.S., Salm, S.N., Xiong, X., Kamb, A., Wesche, H., Marshall, L., Cutler, G., Wang, X., Zavadil, J., Moscatelli, D. & Wilson, E.L., Molecular Signatures of the Primitive Prostate Stem Cell Niche Reveal Novel Mesenchymal-Epithelial Signaling Pathways (Prostate Stem Cell Niche Genes). *PLoS ONE,* 5(9), p. e13024.

Bonkhoff, H., Stein, U. & Remberger, K., 1994. THE PROLIFERATIVE FUNCTION OF BASAL CELLS IN THE NORMAL AND HYPERPLASTIC HUMAN PROSTATE. *Prostate,* 24(3), pp. 114-118.

Bonnet, D. & Dick, J.E., 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine,* 3(7), pp. 730-737.

Boorjian, S.A., Karnes, R.J., Viterbo, R., Rangel, L.J., Bergstralh, E.J., Horwitz, E.M., Blute, M.L. & Buyyounouski, M.K., 2011. Long-Term Survival After Radical Prostatectomy Versus External-Beam Radiotherapy for Patients With High-Risk Prostate Cancer. *Cancer,* 117(13), pp. 2883-2891.

Booth, H.A.F. & Holland, P.W.H., 2004. Eleven daughters of NANOG. *Genomics,* 84(2), pp. 229-238.

Bostwick, D.G. & Brawer, M.K., 1987. PROSTATIC INTRAEPITHELIAL NEOPLASIA AND EARLY INVASION IN PROSTATE-CANCER. *Cancer,* 59(4), pp. 788-794.

Bowen, C., Bubendorf, L., Voeller, H.J., Slack, R., Willi, N., Sauter, G., Gasser, T.C., Koivisto, P., Lack, E.E., Kononen, J., Kallioniemi, O.P. & Gelmann, E.P., 2000. Loss of NKX3.1 expression in human prostate cancers correlates with tumor progression. *Cancer Research,* 60(21), pp. 6111-6115.

Braun, M., Menon, R., Nikolov, P., Kirsten, R., Petersen, K., Schilling, D., Schott, C., Gundisch, S., Fend, F., Becker, K.F. & Perner, S., 2011. The HOPE fixation technique - a promising alternative to common prostate cancer biobanking approaches. *Bmc Cancer,* 11.

Braun, M., Stomper, J., Kirsten, R., Shaikhibrahim, Z., Vogel, W., Bohm, D., Wernert, N., Kristiansen, G. & Perner, S., 2013. Landscape of chromosome number changes in prostate cancer progression. *World Journal of Urology,* 31(6), pp. 1489-1495.

Brimo, F. & Epstein, J.I., 2012. Immunohistochemical pitfalls in prostate pathology. *Human Pathology,* 43(3).

Brimo, F., Vollmer, R.T., Friszt, M., Corcos, J. & Bismar, T.A., 2010. Syndecan-1 expression in prostate cancer and its value as biomarker for disease progression. *Bju International,* 106(3), pp. 418-423.

Brocker, C., Lassen, N., Estey, T., Pappa, A., Cantore, M., Orlova, V.V., Chavakis, T., Kavanagh, K.L., Oppermann, U. & Vasiliou, V., 2010. Aldehyde Dehydrogenase 7A1 (ALDH7A1) Is a Novel Enzyme Involved in Cellular Defense against Hyperosmotic Stress. *Journal of Biological Chemistry,* 285(24), pp. 18452-18463.

Brown, M.D., Gilmore, P.E., Hart, C.A., Samuel, J.D., Ramani, V.A.C., George, N.J. & Clarke, N.W., 2007. Characterization of benign and malignant prostate epithelial Hoechst 33342 side populations. *Prostate,* 67(13), pp. 1384-1396.

Bruchova, H., Vasikova, A., Merkerova, M., Milcova, A., Topinka, J., Balascak, I., Pastorkova, A., Sram, R.J. & Brdicka, R., 2010. Effect of Maternal Tobacco Smoke Exposure on the Placental Transcriptome. *Placenta,* 31(3), pp. 186-191.

Bubendorf, L., Kononen, J., Koivisto, P., Schraml, P., Moch, H., Gasser, T.C., Willi, N., Mihatsch, M.J., Sauter, G. & Kallioniemi, O.P., 1999. Survey of gene amplifications during prostate cancer progression by high throughput fluorescence in situ hybridization on tissue microarrays. *Cancer Research,* 59(4), pp. 803-806.

Burden, H.P., Holmes, C.H., Persad, R. & Whittington, K., 2006. Prostasomes - their effects on human male reproduction and fertility. *Human Reproduction Update,* 12(3), pp. 283-292.

Burger, P.E., Gupta, R., Xiong, X., Ontiveros, C.S., Salm, S.N., Moscatelli, D. & Wilson, E.L., 2009. High Aldehyde Dehydrogenase Activity: A Novel Functional Marker of Murine Prostate Stem/Progenitor Cells. *Stem Cells,* 27(9).

Camp, R.L., Neumeister, V. & Rimm, D.L., 2008. A Decade of Tissue Microarrays: Progress in the Discovery and Validation of Cancer Biomarkers. *Journal of Clinical Oncology,* 26(34), pp. 5630-5637.

Cao, Q., Mani, R.S., Ateeq, B., Dhanasekaran, S.M., Asangani, I.A., Prensner, J.R., Kim, J.H., Brenner, J.C., Jing, X.J., Cao, X.H., Wang, R., Li, Y., Dahiya, A., Wang, L., Pandhi, M., Lonigro, R.J., Wu, Y.M., Tomlins, S.A., Palanisamy, N., Qin, Z.H., Yu, J.D., Maher, C.A., Varambally, S. & Chinnaiyan, A.M., 2011. Coordinated Regulation of Polycomb Group Complexes through microRNAs in Cancer. *Cancer Cell,* 20(2), pp. 187-199.

Carreira, S., Romanel, A., Goodall, J., Grist, E., Ferraldeschi, R., Miranda, S., Prandi, D., Lorente, D., Frenel, J.S., Pezaro, C., Omlin, A., Rodrigues, D.N., Flohr, P., Tunariu, N., de Bono, J.S., Demichelis, F. & Attard, G., 2014. Tumor clone dynamics in lethal prostate cancer. *Science Translational Medicine,* 6(254).

Carter, M.G., Stagg, C.A., Falco, G., Yoshikawa, T., Bassey, U.C., Aiba, K., Sharova, L.V., Shaik, N. & Ko, M.S.H., 2008. An in situ hybridzation-based screen for heterogeneously expressed genes in mouse ES cells. *Gene Expression Patterns,* 8(3).

Carver, B.S., Chapinski, C., Wongvipat, J., Hieronymus, H., Chen, Y., Chandarlapaty, S., Arora, V.K., Le, C., Koutcher, J., Scher, H., Scardino, P.T., Rosen, N. & Sawyers, C.L., 2011. Reciprocal Feedback Regulation of PI3K and Androgen Receptor Signaling in PTEN-Deficient Prostate Cancer. *Cancer Cell,* 19(5), pp. 575-586.

Castellon, E.A., Valenzuela, R., Lillo, J., Castillo, V., Contreras, H.R., Gallegos, I., Mercado, A. & Huidobro, C., 2012. Molecular signature of cancer stem cells isolated from prostate carcinoma and expression of stem markers in different Gleason grades and metastasis. *Biological Research,* 45(3), pp. 297-305.

Catalona, W.J., Smith, D.S., Ratliff, T.L., Dodds, K.M., Coplen, D.E., Yuan, J.J.J., Petros, J.A. & Andriole, G.L., 1991. MEASUREMENT OF PROSTATE-SPECIFIC ANTIGEN IN SERUM AS A SCREENING-TEST FOR PROSTATE-CANCER. *New England Journal of Medicine,* 324(17), pp. 1156-1161.

Celia-Terrassa, T., Meca-Cortes, O., Mateo, F., de Paz, A.M., Rubio, N., Arnal-Estape, A., Ell, B.J., Bermudo, R., Diaz, A., Guerra-Rebollo, M., Lozano, J.J., Estaras, C., Ulloa, C., Alvarez-Simon, D., Mila, J., Vilella, R., Paciucci, R., Martinez-Balbas, M., de Herreros, A.G., Gomis, R.R., Kang, Y.B., Blanco, J., Fernandez, P.L. & Thomson, T.M., 2012. Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. *Journal of Clinical Investigation,* 122(5), pp. 1849-1868.

Chaffer, C.L., Brueckmann, I., Scheel, C., Kaestli, A.J., Wiggins, P.A., Rodrigues, L.O., Brooks, M., Reinhardt, F., Su, Y., Polyak, K., Arendt, L.M., Kuperwasser, C., Bierie, B. & Weinberg, R.A., 2011. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proceedings of the National Academy of Sciences of the United States of America,* 108(19), pp. 7950-7955.

Chan, C.-L., Wong, J.W.Y., Wong, C.-P., Chan, M.K.L. & Fong, W.-P., 2011. Human antiquitin: Structural and functional studies. *Chemico-Biological Interactions,* 191(1-3), pp. 165-170.

Chan, J.M., Stampfer, M.J., Ma, J., Gann, P., Gaziano, J.M., Pollak, M. & Giovannucci, E., 2002. Insulin-like growth factor-I (IGF-I) and IGF binding protein-3 as predictors of advanced-stage prostate cancer. *Journal of the National Cancer Institute,* 94(14), pp. 1099-1106.

Chen, D., Adenekan, B., Chen, L., Vaughan, E.D., Gerald, W., Feng, Z.D. & Knudsen, B.S., 2004a. Syndecan-1 expression in locally invasive and metastatic prostate cancer. *Urology,* 63(2), pp. 402-407.

Chen, M.L., Xu, P.Z., Peng, X.D., Chen, W.S., Guzman, G., Yang, X.M., Di Cristofano, A., Pandolfi, P.P. & Hay, N., 2006. The deficiency of Akt1 is sufficient to suppress tumor development in Pten(+/-) mice. *Genes & Development,* 20(12), pp. 1569-1574.

Chen, Y., Chi, P., Rockowitz, S., Iaquinta, P.J., Shamu, T., Shukla, S., Gao, D., Sirota, I., Carver, B.S., Wongvipat, J., Scher, H.I., Zheng, D.Y. & Sawyers, C.L., 2013. ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. *Nature Medicine,* 19(8), pp. 1023-+.

Chen, Z.B., Trotman, L.C., Shaffer, D., Lin, H.K., Dotan, Z.A., Niki, M., Koutcher, J.A., Scher, H.I., Ludwig, T., Gerald, W., Cordon-Cardo, C. & Pandolfi, P.P., 2005. Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature,* 436(7051), pp. 725-730.

Cheng, J., Yang, K., Zhang, Q., Yu, Y., Meng, Q., Mo, N., Zhou, Y., Yi, X., Ma, C., Lei, A. & Liu, Y., 2016. The role of mesenchymal stem cells in promoting the transformation of androgen-dependent human prostate cancer cells into androgen-independent manner. *Scientific Reports,* 6.

Cheng, L., Montironi, R., Bostwick, D.G., Lopez-Beltran, A. & Berney, D.M., 2012. Staging of prostate cancer. *Histopathology,* 60(1).

Cheng, L., Song, S.Y., Pretlow, T.G., Abdul-Karim, F.W., Kung, H.J., Dawson, D.V., Park, W.S., Moon, Y.W., Tsai, M.L., Linehan, W.M., Emmert-Buck, M.R., Liotta, L.A. & Zhuang, Z.P., 1998. Evidence of independent origin of multiple tumors from patients with prostate cancer. *Journal of the National Cancer Institute,* 90(3), pp. 233-237.

Chien, C.S., Wang, M.L., Chu, P.Y., Chang, Y.L., Liu, W.H., Yu, C.C., Lan, Y.T., Huang, P.I., Lee, Y.Y., Chen, Y.W., Lo, W.L. & Chiou, S.H., 2015. Lin28B/Let-7 Regulates Expression of Oct4 and Sox2 and Reprograms Oral Squamous Cell Carcinoma Cells to a Stem-like State. *Cancer Research,* 75(12), pp. 2553-2565.

Chilosi, M., Adami, F., Lestani, M., Montagna, L., Cimarosto, L., Semenzato, G., Pizzolo, G. & Menestrina, F., 1999. CD138/syndecan-1: A useful immunohistochemical marker of normal and neoplastic plasma cells on routine trephine bone marrow biopsies. *Modern Pathology,* 12(12), pp. 1101-1106.

Cho, Y.M., Kim, Y.S., Kang, M.J., Farrar, W.L. & Hurt, E.M., 2012. Long-term recovery of irradiated prostate cancer increases cancer stem cells. *The Prostate,* 72(16), pp. 1746-56.

Choi, N., Zhang, B., Zhang, L., Ittmann, M. & Xin, L., 2012. Adult Murine Prostate Basal and Luminal Cells Are Self-Sustained Lineages that Can Both Serve as Targets for Prostate Cancer Initiation. *Cancer Cell,* 21(2), pp. 253-265.

Christoffersen, N.R., Shalgi, R., Frankel, L.B., Leucci, E., Lees, M., Klausen, M., Pilpel, Y., Nielsen, F.C., Oren, M. & Lund, A.H., 2010. p53-independent upregulation of miR-34a during oncogene-induced senescence represses MYC. *Cell Death and Differentiation,* 17(2), pp. 236-245.

Cole, A.I., Morgan, T.M., Spratt, D.E., Palapattu, G.S., He, C., Tomlins, S.A., Weizer, A.Z., Feng, F.Y., Wu, A., Siddiqui, J., Chinnaiyan, A.M., Montgomery, J.S., Kunju, L.P., Miller, D.C., Hollenbeck, B.K., Wei, J.T. & Mehra, R., 2016. Prognostic Value of Percent Gleason Grade 4 at Prostate Biopsy in Predicting Prostatectomy Pathology and Recurrence. *Journal of Urology,* 196(2), pp. 405-411.

Coller, H.A., Grandori, C., Tamayo, P., Colbert, T., Lander, E.S., Eisenman, R.N. & Golub, T.R., 2000. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proceedings of the National Academy of Sciences of the United States of America,* 97(7), pp. 3260-3265.

Collins, A.T., Berry, P.A., Hyde, C., Stower, M.J. & Maitland, N.J., 2005. Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Research,* 65(23), pp. 10946-10951.

Contreras, H.R., Ledezma, R.A., Vergara, J., Cifuentes, F., Barra, C., Cabello, P., Gallegos, I., Morales, B., Huidobro, C. & Castellon, E.A., 2010. The expression of syndecan-1 and-2 is associated with Gleason score and epithelial-mesenchymal transition markers, E-cadherin and beta-catenin, in prostate cancer. *Urologic Oncology-Seminars and Original Investigations,* 28(5).

Costello, L.C. & Franklin, R.B., 1998. Novel role of zinc in the regulation of prostate citrate metabolism and its implications in prostate cancer. *Prostate,* 35(4), pp. 285-296.

Crawford, E.D. & Abrahamsson, P.A., 2008. PSA-based screening for prostate cancer: How does it compare with other cancer screening tests? *European Urology,* 54(2), pp. 262-273.

Crea, F., Duhagon Serrat, M.A., Hurt, E.M., Thomas, S.B., Danesi, R. & Farrar, W.L., 2011a. BMI1 silencing enhances docetaxel activity and impairs antioxidant response in prostate cancer. *International Journal of Cancer,* 128(8), pp. 1946-1954.

Crea, F., Hurt, E.M., Mathews, L.A., Cabarcas, S.M., Sun, L., Marquez, V.E., Danesi, R. & Farrar, W.L., 2011b. Pharmacologic disruption of Polycomb Repressive Complex 2 inhibits tumorigenicity and tumor progression in prostate cancer. *Molecular Cancer,* 10.

Culig, Z., Hobisch, A., Cronauer, M.V., Cato, A.C.B., Hittmair, A., Radmayr, C., Eberle, J., Bartsch, G. & Klocker, H., 1993. MUTANT ANDROGEN RECEPTOR DETECTED IN AN ADVANCED-STAGE PROSTATIC-CARCINOMA IS ACTIVATED BY ADRENAL ANDROGENS AND PROGESTERONE. *Molecular Endocrinology,* 7(12), pp. 1541-1550.

D'Amico, A.V., Whittington, R., Malkowicz, S.B., Cote, K., Loffredo, M., Schultz, D., Chen, M.H., Tomaszewski, J.E., Renshaw, A.A., Wein, A. & Richie, J.P., 2002. Biochemical outcome after radical prostatectomy or external beam radiation therapy for patients with clinically localized prostate carcinoma in the prostate specific antigen era. *Cancer,* 95(2), pp. 281-286.

Dang, C.V., 2012. MYC on the Path to Cancer. *Cell,* 149(1), pp. 22-35.

Das, S., Jena, S. & Levasseur, D.N., 2011. Alternative Splicing Produces Nanog Protein Variants with Different Capacities for Self-renewal and Pluripotency in Embryonic Stem Cells. *Journal of Biological Chemistry,* 286(49), pp. 42690-42703.

Dave, B., Landis, M.D., Dobrolecki, L.E., Wu, M.-F., Zhang, X., Westbrook, T.F., Hilsenbeck, S.G., Liu, D., Lewis, M.T., Tweardy, D.J. & Chang, J.C., 2012. Selective Small Molecule Stat3 Inhibitor Reduces Breast Cancer Tumor-Initiating Cells and Improves Recurrence Free Survival in a Human-Xenograft Model. *Plos One,* 7(8).

Davis, J.N. & Day, M.L., 2002. The convergence of hormone regulation and cell cycle in prostate physiology and prostate tumorigenesis. *Molecular Biotechnology,* 22(2), pp. 129-138.

De Marzo, A.M., Bradshaw, C., Sauvageot, J., Epstein, J.I. & Miller, G.J., 1998. CD44 and CD44v6 downregulation in clinical prostatic carcinoma: Relation to Gleason Grade and cytoarchitecture. *Prostate,* 34(3), pp. 162-168.

De Marzo, A.M., Marchi, V.L., Epstein, J.I. & Nelson, W.G., 1999. Proliferative inflammatory atrophy of the prostate - Implications for prostatic carcinogenesis. *American Journal of Pathology,* 155(6), pp. 1985-1992.

De Marzo, A.M., Meeker, A.K., Zha, S., Luo, J., Nakayama, M., Platz, E.A., Isaacs, W.B. & Nelson, W.G., 2003. Human prostate cancer precursors and pathobiology. *Urology,* 62(5A), pp. 55-62.

de Resende, M.F., Chinen, L.T.D., Vieira, S., Jampietro, J., da Fonseca, F.P., Vassallo, J., Campos, L.C., Guimares, G.C., Soares, F.A. & Rocha, R.M., 2013. Prognostication of OCT4 isoform expression in prostate cancer. *Tumor Biology,* 34(5), pp. 2665-2673.

De Vos, J., Hose, D., Reme, T., Tarte, K., Moreaux, J., Mahtouk, K., Jourdan, M., Goldschmidt, H., Rossi, J.F., Cremer, F.W. & Klein, B., 2006. Microarray-based understanding of normal and malignant plasma cells. *Immunological Reviews,* 210, pp. 86-104.

DeMarzo, A.M., Nelson, W.G., Isaacs, W.B. & Epstein, J.I., 2003. Pathological and molecular aspects of prostate cancer. *Lancet,* 361(9361), pp. 955-964.

Deng, F.M., Donin, N.M., Benito, R.P., Melamed, J., Le Nobin, J., Zhou, M., Ma, S.S., Wang, J.H. & Lepor, H., 2016. Size-adjusted Quantitative Gleason Score as a Predictor of Biochemical Recurrence after Radical Prostatectomy. *European Urology,* 70(2), pp. 248-253.

Dent, R., Trudeau, M., Pritchard, K.I., Hanna, W.M., Kahn, H.K., Sawka, C.A., Lickley, L.A., Rawlinson, E., Sun, P. & Narod, S.A., 2007. Triple-negative breast cancer: Clinical features and patterns of recurrence. *Clinical Cancer Research,* 13(15), pp. 4429-4434.

Di Cristofano, A., De Acetis, M., Koff, A., Cordon-Cardo, C. & Pandolfi, P.P., 2001. Pten and p27(KIP1) cooperate in prostate cancer tumor suppression in the mouse. *Nature Genetics,* 27(2), pp. 222-224.

diSantAgnese, P.A. & Cockett, A.T.K., 1996. Neuroendocrine differentiation in prostatic malignancy. *Cancer,* 78(2), pp. 357-361.

Doganavsargil, B., Simsir, A., Boyacioglu, H., Cal, C. & Hekimgil, M., 2006. A comparison of p21 and p27 immunoexpression in benign glands, prostatic intraepithelial neoplasia and prostate adenocarcinoma. *Bju International,* 97(3), pp. 644-648.

Doherty, R.E., Haywood-Small, S.L., Sisley, K. & Cross, N.A., 2011. Aldehyde dehydrogenase activity selects for the holoclone phenotype in prostate cancer cells. *Biochemical and Biophysical Research Communications,* 414(4), pp. 801-807.

Domingo-Domenech, J., Vidal, S.J., Rodriguez-Bravo, V., Castillo-Martin, M., Quinn, S.A., Rodriguez-Barrueco, R., Bonal, D.M., Charytonowicz, E., Gladoun, N., de la Iglesia-Vicente, J., Petrylak, D.P., Benson, M.C., Silva, J.M. & Cordon-Cardo, C., 2012. Suppression of Acquired Docetaxel Resistance in Prostate Cancer through Depletion of Notch- and Hedgehog-Dependent Tumor-Initiating Cells. *Cancer Cell,* 22(3), pp. 373-388.

Donovan, M.J., Hamann, S., Clayton, M., Khan, F.M., Sapir, M., Bayer-Zubek, V., Fernandez, G., Mesa-Tejada, R., Teverovskiy, M., Reuter, V.E., Scardino, P.T. & Cordon-Cardo, C., 2008. Systems pathology approach for the prediction of prostate cancer progression after radical prostatectomy. *Journal of Clinical Oncology,* 26(24), pp. 3923-3929.

Duhagon, M.A., Hurt, E.M., Sotelo-Silveira, J.R., Zhang, X.H. & Farrar, W.L., 2010. Genomic profiling of tumor initiating prostatospheres. *Bmc Genomics,* 11, p. 16.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B.R., Landt, S.G., Lee, B.K., Pauli, F., Rosenbloom, K.R., Sabo, P., Safi, A., Sanyal, A., Shoresh, N., Simon, J.M., Song, L., Trinklein, N.D., Altshuler, R.C., Birney, E., Brown, J.B., Cheng, C., Djebali, S., Dong, X.J., Ernst, J., Furey, T.S., Gerstein, M., Giardine, B., Greven, M., Hardison, R.C., Harris, R.S., Herrero, J., Hoffman, M.M., Iyer, S., Kellis, M., Kheradpour, P., Lassmann, T., Li, Q.H., Lin, X., Marinov, G.K., Merkel, A., Mortazavi, A., Parker, S.C.J., Reddy, T.E., Rozowsky, J., Schlesinger, F., Thurman, R.E., Wang, J., Ward, L.D., Whitfield, T.W., Wilder, S.P., Wu, W., Xi, H.L.S., Yip, K.Y., Zhuang, J.L., Bernstein, B.E., Green, E.D., Gunter, C., Snyder, M., Pazin, M.J., Lowdon, R.F., Dillon, L.A.L., Adams, L.B., Kelly, C.J., Zhang, J., Wexler, J.R., Good, P.J., Feingold, E.A., Crawford, G.E., Dekker, J., Elnitski, L., Farnham, P.J., Giddings, M.C., Gingeras, T.R., Guigo, R., Hubbard, T.J., Kent, W.J., Lieb, J.D., Margulies, E.H., Myers, R.M., Stamatoyannopoulos, J.A., Tenenbaum, S.A., Weng, Z.P., White, K.P., Wold, B., Yu, Y., Wrobel, J., Risk, B.A., Gunawardena, H.P., Kuiper, H.C., Maier, C.W., Xie, L., Chen, X., Mikkelsen, T.S., Gillespie, S., *et al.*, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489(7414), pp. 57-74.

Dutta, A., Le Magnen, C., Mitrofanova, A., Ouyang, X.S., Califano, A. & Abate-Shen, C., 2016. Identification of an NKX3.1-G9a-UTY transcriptional regulatory network that controls prostate differentiation. *Science,* 352(6293), pp. 1576-1580.

Dzinic, S.H., Chen, K., Thakur, A., Kaplun, A., Bonfil, R.D., Li, X.H., Liu, J.S., Bernardo, M.M., Saliganan, A., Back, J.B., Yano, H., Schalk, D.L., Tomaszewski, E.N., Beydoun, A.S., Dyson, G., Mujagic, A., Krass, D., Dean, I., Mi, Q.S., Heath, E., Sakr, W., Lum, L.G. & Sheng, S.J., 2014. No Maspin expression in prostate tumor elicits host anti-tumor immunity. *Oncotarget,* 5(22), pp. 11225-11236.

Ebelt, K., Babaryka, G., Frankenberger, B., Stief, C.G., Eisenmenger, W., Kirchner, T., Schendel, D.J. & Noessner, E., 2009. Prostate cancer lesions are surrounded by FOXP3(+), PD-1(+) and B7-H1(+) lymphocyte clusters. *European Journal of Cancer,* 45(9), pp. 1664-1672.

Ecke, T.H., Schlechte, H.H., Schiemenz, K., Sachs, M.D., Lenk, S.V., Rudolph, B.D. & Loening, S.A., 2010. TP53 Gene Mutations in Prostate Cancer Progression. *Anticancer Research,* 30(5), pp. 1579-1586.

Eckel-Passow, J.E., Lohse, C.M., Sheinin, Y., Crispen, P.L., Krco, C.J. & Kwon, E.D., 2010. Tissue microarrays: one size does not fit all. *Diagnostic Pathology,* 5.

Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D., 1998. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America,* 95(25), pp. 14863-14868.

El-Alfy, M., Pelletier, G., Hermo, L.S. & Labrie, F., 2000. Unique features of the basal cells of human prostate epithelium. *Microscopy Research and Technique,* 51(5), pp. 436-446.

Ellwood-Yen, K., Graeber, T.G., Wongvipat, J., Iruela-Arispe, M.L., Zhang, J.F., Matusik, R., Thomas, G.V. & Sawyers, C.L., 2003. Myc-driven murine prostate cancer shares molecular features with human prostate tumors. *Cancer Cell,* 4(3), pp. 223-238.

Engelmann, K., Shen, H. & Finn, O.J., 2008. MCF7 side population cells with characteristics of cancer stem/progenitor cells express the tumor antigen MUC1. *Cancer Research,* 68(7), pp. 2419-2426.

Epstein, J.I., 2004. Diagnosis and reporting of limited adenocarcinoma of the prostate on needle biopsy. *Modern Pathology,* 17(3), pp. 307-315.

Epstein, J.I., 2009. Precursor lesions to prostatic adenocarcinoma. *Virchows Archiv,* 454(1), pp. 1-16.

Epstein, J.I., Egevad, L., Amin, M.B., Delahunt, B., Srigley, J.R., Humphrey, P.A. & Grading, C., 2016a. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma Definition of Grading Patterns and Proposal for a New Grading System. *American Journal of Surgical Pathology,* 40(2), pp. 244-252.

Epstein, J.I., Pizov, G. & Walsh, P.C., 1993. CORRELATION OF PATHOLOGICAL FINDINGS WITH PROGRESSION AFTER RADICAL RETROPUBIC PROSTATECTOMY. *Cancer,* 71(11), pp. 3582-3593.

Epstein, J.I., Zelefsky, M.J., Sjoberg, D.D., Nelson, J.B., Egevad, L., Magi-Galluzzi, C., Vickers, A.J., Parwani, A.V., Reuter, V.E., Fine, S.W., Eastham, J.A., Wiklund, P., Han, M., Reddy, C.A., Ciezki, J.P., Nyberg, T. & Klein, E.A., 2016b. A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. *European Urology,* 69(3), pp. 428-435.

National Institute for Health and Care Excellence (NICE), 2015. Suspected cancer: recognition and referral. NICE guideline (NG12).

Falco, G., Lee, S.L., Stanghellini, I., Bassey, U.C., Hamatani, T. & Ko, M.S.H., 2007. Zscan4: A novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Developmental Biology,* 307(2), pp. 539-550.

Feneley, M.R., Young, M.P.A., Chinyama, C., Kirby, R.S. & Parkinson, M.C., 1996. Ki-67 expression in early prostate cancer and associated pathological lesions. *Journal of Clinical Pathology,* 49(9), pp. 741-748.

Fenic, I., Franke, F., Failing, K., Steger, K. & Woenckhaus, J., 2004. Expression of PTEN in malignant and non-malignant human prostate tissues: comparison with p27 protein expression. *Journal of Pathology,* 203(1), pp. 559-566.

Fine, S.W. & Reuter, V.E., 2012. Anatomy of the prostate revisited: implications for prostate biopsy and zonal origins of prostate cancer. *Histopathology,* 60(1).

Finones, R.R., Yeargin, J., Lee, M., Kaur, A.P., Cheng, C., Sun, P., Wu, C., Nguyen, C., Wang-Rodriguez, J., Meyer, A.N., Baird, S.M., Donoghue, D.J. & Haas, M., 2013. Early Human Prostate Adenocarcinomas Harbor Androgen-Independent Cancer Cells. *PLoS One,* 8(9), p. e74438.

Flavin, R., Pettersson, A., Hendrickson, W.K., Fiorentino, M., Finn, S., Kunz, L., Judson, G.L., Lis, R., Bailey, D., Fiore, C., Nuttall, E., Martin, N.E., Stack, E., Penney, K.L., Rider, J.R., Sinnott, J., Sweeney, C., Sesso, H.D., Fall, K., Giovannucci, E., Kantoff, P., Stampfer, M., Loda, M. & Mucci, L.A., 2014. SPINK1 Protein Expression and Prostate Cancer Progression. *Clinical Cancer Research,* 20(18), pp. 4904-4911.

Fleischmann, A., Schlomm, T., Kollermann, J., Sekulic, N., Huland, H., Mirlacher, M., Sauter, G., Simon, R. & Erbersdobler, A., 2009. Immunological Microenvironment in Prostate Cancer: High Mast Cell Densities Are Associated With Favorable Tumor Characteristics and Good Prognosis. *Prostate,* 69(9), pp. 976-981.

Foster, C.S., Dodson, A., Karavana, V., Smith, P.H. & Ke, Y., 2002. Prostatic stem cells. *Journal of Pathology,* 197(4), pp. 551-565.

Foster, C.S., Falconer, A., Dodson, A.R., Norman, A.R., Dennis, N., Fletcher, A., Southgate, C., Dowe, A., Dearnaley, D., Jhavar, S., Eeles, R., Feber, A. & Cooper, C.S., Transcription factor E2F3 overexpressed in prostate cancer independently predicts clinical outcome. *Oncogene,* 23(35), pp. 5871-5879.

Fu, Q.B., Quan, Y., Wang, W.K., Mei, T., Wu, J.W., Li, J., Yang, G., Ren, X.T., Xue, J.M. & Wang, Y.G., 2012. Response of cancer stem-like cells and non-stem cancer cells to proton and gamma-ray irradiation. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms,* 286, pp. 346-350.

Fujii, T., Shimada, K., Tatsumi, Y., Fujimoto, K. & Konishi, N., 2015. Syndecan-1 responsive microRNA-126 and 149 regulate cell proliferation in prostate cancer. *Biochemical and biophysical research communications,* 456(1), pp. 183-9.

Gao, A.C., Lou, W., Dong, J.T. & Isaacs, J.T., 1997. CD44 is a metastasis suppressor gene for prostatic cancer located on human chromosome 11p13. *Cancer Research,* 57(5), pp. 846-849.

Gao, H., Ouyang, X.S., Banach-Petrosky, W.A., Shen, M.M. & Abate-Shen, C., 2006. Emergence of androgen independence at early stages of prostate cancer progression in Nkx3.1; Pten mice. *Cancer Research,* 66(16), pp. 7929-7933.

Geng, C., He, B., Xu, L.M., Barbieri, C.E., Eedunuri, V.K., Chew, S.A., Zimmermann, M., Bond, R., Shou, J., Li, C., Blattner, M., Lonard, D.M., Demichelis, F., Coarfa, C., Rubin, M.A., Zhou, P., O'Malley, B.W. & Mitsiades, N., 2013. Prostate cancer-associated mutations in speckle-type POZ protein (SPOP) regulate steroid receptor coactivator 3 protein turnover. *Proceedings of the National Academy of Sciences of the United States of America,* 110(17), pp. 6997-7002.

Gerdes, J., Lemke, H., Baisch, H., Wacker, H.H., Schwab, U. & Stein, H., 1984. CELL-CYCLE ANALYSIS OF A CELL PROLIFERATION-ASSOCIATED HUMAN NUCLEAR ANTIGEN DEFINED BY THE MONOCLONAL-ANTIBODY KI-67. *Journal of Immunology,* 133(4), pp. 1710-1715.

Germann, M., Wetterwald, A., Guzman-Ramirez, N., van der Pluijm, G., Culig, Z., Cecchini, M.G., Williams, E.D. & Thalmann, G.N., 2012. Stem-Like Cells with Luminal Progenitor Phenotype Survive Castration in Human Prostate Cancer. *Stem Cells,* 30(6), pp. 1076-1086.

Gharbaran, R., 2015. Advances in the molecular functions of syndecan-1 (SDC1/CD138) in the pathogenesis of malignancies. *Critical Reviews in Oncology Hematology,* 94(1), pp. 1-17.

Gleason, D.F. & Mellinge.Gt, 1974. PREDICTION OF PROGNOSIS FOR PROSTATIC ADENOCARCINOMA BY COMBINED HISTOLOGICAL GRADING AND CLINICAL STAGING. *Journal of Urology,* 111(1), pp. 58-64.

Glinsky, G.V., Berezovska, O. & Glinskii, A.B., 2005. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. *Journal of Clinical Investigation,* 115(6).

Goldstein, A.S., Huang, J., Guo, C., Garraway, I.P. & Witte, O.N., 2010. Identification of a Cell of Origin for Human Prostate Cancer. *Science,* 329(5991), pp. 568-571.

Goltz, D., Montani, M., Braun, M., Perner, S., Wernert, N., Jung, K., Dietel, M., Stephan, C. & Kristiansen, G., 2015. Prognostic relevance of proliferation markers (Ki-67, PHH3) within the cross-relation of ERG translocation and androgen receptor expression in prostate cancer. *Pathology,* 47(7), pp. 629-636.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D. & Lander, E.S., 1999. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science,* 286(5439), pp. 531-537.

Gordetsky, J. & Epstein, J., 2016. Grading of prostatic adenocarcinoma: current state and prognostic implications. *Diagnostic Pathology,* 11.

Grasso, C.S., Wu, Y.-M., Robinson, D.R., Cao, X., Dhanasekaran, S.M., Khan, A.P., Quist, M.J., Jing, X., Lonigro, R.J., Brenner, J.C., Asangani, I.A., Ateeq, B., Chun, S.Y., Siddiqui, J., Sam, L., Anstett, M., Mehra, R., Prensner, J.R., Palanisamy, N., Ryslik, G.A., Vandin, F., Raphael, B.J., Kunju, L.P., Rhodes, D.R., Pienta, K.J., Chinnaiyan, A.M. & Tomlins, S.A., 2012. The mutational landscape of lethal castration-resistant prostate cancer. *Nature,* 487(7406), pp. 239-243.

Gravdal, K., Halvorsen, O.J., Haukaas, S.A. & Akslen, L.A., 2009. Proliferation of Immature Tumor Vessels Is a Novel Marker of Clinical Progression in Prostate Cancer. *Cancer Research,* 69(11).

Groner, A.C., Cato, L., de Tribolet-Hardy, J., Bernasocchi, T., Janouskova, H., Melchers, D., Houtman, R., Cato, A.C.B., Tschopp, P., Gu, L., Corsinotti, A., Zhong, Q., Fankhauser, C., Fritz, C., Poyet, C., Wagner, U., Guo, T., Aebersold, R., Garraway, L.A., Wild, P.J., Theurillat, J.P. & Brown, M., 2016. TRIM24 Is an Oncogenic Transcriptional Activator in Prostate Cancer. *Cancer Cell,* 29(6), pp. 846-858.

Grupp, K., Diebel, F., Sirma, H., Simon, R., Breitmeyer, K., Steurer, S., Hube-Magg, C., Prien, K., Pham, T., Weigand, P., Michl, U., Heinzer, H., Kluth, M., Minner, S., Tsourlakis, M.C., Izbicki, J.R., Sauter, G., Schlomm, T. & Wilczak, W., 2013. SPINK1 Expression is Tightly Linkedto 6q15-and 5q21-Deleted ERG-Fusion Negative Prostate Cancers But Unrelated to PSA Recurrence. *Prostate,* 73(15), pp. 1690-1698.

Gu, G.Y., Yuan, J.L., Wils, M. & Kasper, S., 2007. Prostate cancer cells with stem cell characteristics reconstitute the original human tumor in vivo. *Cancer Research,* 67(10), pp. 4807-4815.

Gupta, P.B., Fillmore, C.M., Jiang, G.Z., Shapira, S.D., Tao, K., Kuperwasser, C. & Lander, E.S., 2011. Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells. *Cell,* 146(4), pp. 633-644.

Gurel, B., Iwata, T., Koh, C.M., Jenkins, R.B., Lan, F.S., Van Dang, C., Hicks, J.L., Morgan, J., Cornish, T.C., Sutcliffe, S., Isaacs, W.B., Luo, J. & De Marzo, A.M., 2008. Nuclear MYC protein overexpression is an early alteration in human prostate carcinogenesis. *Modern Pathology,* 21(9), pp. 1156-1167.

Hamilton, W. & Sharp, D., 2004. Symptomatic diagnosis of prostate cancer in primary care: a structured review. *British Journal of General Practice,* 54(505), pp. 617-621.

Hara, T., Miyazaki, J., Araki, H., Yamaoka, M., Kanzaki, N., Kusaka, M. & Miyamoto, M., 2003. Novel mutations of androgen receptor: A possible mechanism of bicalutamide withdrawal syndrome. *Cancer Research,* 63(1), pp. 149-153.

Harsha, H.C., Kandasamy, K., Ranganathan, P., Rani, S., Ramabadran, S., Gollapudi, S., Balakrishnan, L., Dwivedi, S.B., Telikicherla, D., Selvan, L.D.N., Goel, R., Mathivanan, S., Marimuthu, A., Kashyap, M., Vizza, R.F., Mayer, R.J., DeCaprio, J.A., Srivastava, S., Hanash, S.M., Hruban, R.H. & Pandey, A., 2009. A Compendium of Potential Biomarkers of Pancreatic Cancer. *Plos Medicine,* 6(4).

Hassan, S., Ferrario, C., Mamo, A. & Basik, M., 2008. Tissue microarrays: emerging standard for biomarker validation. *Current Opinion in Biotechnology,* 19(1).

Hatanaka, Y., Hashizume, K., Nitta, K., Kato, T., Itoh, I. & Tani, Y., 2003. Cytometrical image analysis for immunohistochemical hormone receptor status in breast carcinomas. *Pathology International,* 53(10), pp. 693-699.

Hattrup, C.L. & Gendler, S.J., 2008. Structure and function of the cell surface (tethered) mucins. *Annual Review of Physiology.* pp. 431-457.

Hawksworth, D., Ravindranath, L., Chen, Y., Furusato, B., Sesterhenn, I.A., McLeod, D.G., Srivastava, S. & Petrovics, G., 2010. Overexpression of C-MYC oncogene in prostate cancer predicts biochemical recurrence. *Prostate Cancer and Prostatic Diseases,* 13(4), pp. 311-315.

Heidenreich, A., Aus, G., Bolla, M., Joniau, S., Matveev, V.B., Schmid, H.P. & Zattoni, F., 2008. EAU guidelines on prostate cancer. *European Urology,* 53(1), pp. 68-80.

Hellsten, R., Johansson, M., Dahlman, A., Sterner, O. & Bjartell, A., 2011. Galiellalactone Inhibits Stem Cell-Like ALDH-Positive Prostate Cancer Cells. *Plos One,* 6(7), p. 9.

Hendriks, J., Gravestein, L.A., Tesselaar, K., van Lier, R.A.W., Schumacher, T.N.M. & Borst, J., 2000. CD27 is required for generation and long-term maintenance of T cell immunity. *Nature Immunology,* 1(5), pp. 433-440.

Hendriksen, P.J.M., Dits, N.F.J., Kokame, K., Veldhoven, A., van Weerden, W.M., Bangma, C.H., Trapman, J. & Jenster, G., 2006. Evolution of the androgen receptor pathway during progression of prostate cancer. *Cancer Research,* 66(10), pp. 5012-5020.

Hernandez, S., Font-Tello, A., Juanpere, N., de Muga, S., Lorenzo, M., Salido, M., Fumado, L., Serrano, L., Cecchini, L., Serrano, S. & Lloreta, J., 2016. Concurrent TMPRSS2-ERG and SLC45A3-ERG Rearrangements Plus PTEN Loss Are Not Found in Low Grade Prostate Cancer and Define an Aggressive Tumor Subset. *Prostate,* 76(9), pp. 854-865.

Hieronymus, H., Schultz, N., Gopalan, A., Carver, B.S., Chang, M.T., Xiao, Y., Heguy, A., Huberman, K., Bernstein, M., Assel, M., Murali, R., Vickers, A., Scardino, P.T., Sander, C., Reuter, V., Taylor, B.S. & Sawyers, C.L., 2014. Copy number alteration burden predicts prostate cancer relapse. *Proceedings of the National Academy of Sciences of the United States of America,* 111(30), pp. 11139-11144.

Hikita, S.T., Kosik, K.S., Clegg, D.O. & Bamdad, C., 2008. MUC1*Mediates the Growth of Human Pluripotent Stem Cells. *Plos One,* 3(10).

Hirano, D., Okada, Y., Minei, S., Takimoto, Y. & Nemoto, N., 2004. Neuroendocrine differentiation in hormone refractory prostate cancer following androgen deprivation therapy. *European Urology,* 45(5), pp. 586-592.

Hirata, T., Amano, T., Nakatake, Y., Amano, M., Piao, Y., Hoang, H.G. & Ko, M.S.H., 2012. Zscan4 transiently reactivates early embryonic genes during the generation of induced pluripotent stem cells. *Scientific Reports,* 2.

Hirschmann-Jax, C., Foster, A.E., Wulf, G.G., Nuchtern, J.G., Jax, T.W., Gobel, U., Goodell, M.A. & Brenner, M.K., 2004. A distinct "side population" of cells with high drug efflux capacity in human tumor cells. *Proceedings of the National Academy of Sciences of the United States of America,* 101(39), pp. 14228-14233.

Holzer, A.K., Varki, N.M., Le, Q.T., Gibson, M.A., Naredi, P. & Howell, S.B., 2006. Expression of the human copper influx transporter 1 in normal and malignant human tissues. *Journal of Histochemistry & Cytochemistry,* 54(9), pp. 1041-1049.

Hong, Q.Y., Sze, C.I., Lin, S.R., Lee, M.H., He, R.Y., Schultz, L., Chang, J.Y., Chen, S.J., Boackle, R.J., Hsu, L.J. & Chang, N.S., 2009. Complement C1q Activates Tumor Suppressor WWOX to Induce Apoptosis in Prostate Cancer Cells. *Plos One,* 4(6).

Hu, J., Liu, Y.-L., Piao, S.-l., Yang, D.-d., Yang, Y.-M. & Cai, L., 2012. Expression patterns of USP22 and potential targets BMI-1, PTEN, p-AKT in non-small-cell lung cancer. *Lung Cancer,* 77(3).

Hu, R., Dunn, T.A., Wei, S.Z., Isharwal, S., Veltri, R.W., Humphreys, E., Han, M., Partin, A.W., Vessella, R.L., Isaacs, W.B., Bova, G.S. & Luo, J., 2009. Ligand-Independent Androgen Receptor Variants Derived from Splicing of Cryptic Exons Signify Hormone-Refractory Prostate Cancer. *Cancer Research,* 69(1), pp. 16-22.

Huang, C.Y., Chang, Y.J., Luo, S.D., Uyanga, B., Lin, F.Y., Tai, C.J. & Huang, M.T., 2016. Maspin mediates the gemcitabine sensitivity of hormone-independent prostate cancer. *Tumor Biology,* 37(3), pp. 4075-4082.

Huggins, C. & Sommer, J.L., 1953. QUANTITATIVE STUDIES OF PROSTATIC SECRETION .3. SIMULTANEOUS MEASUREMENT OF SIZE AND SECRETION OF THE CANINE PROSTATE AND THE INTERACTION OF ANDROGENIC AND ESTROGENIC SUBSTANCES THEREON. *Journal of Experimental Medicine,* 97(5), pp. 663-&.

Humphrey, P.A., 2004. Gleason grading and prognostic factors in carcinoma of the prostate. *Modern Pathology,* 17(3), pp. 292-306.

Hurt, E.M., Chan, K., Duhagon Serrat, M.A., Thomas, S.B., Veenstra, T.D. & Farrar, W.L., 2010. Identification of Vitronectin as an Extrinsic Inducer of Cancer Stem Cell Differentiation and Tumor Formation. *Stem Cells,* 28(3).

Hurt, E.M., Kawasaki, B.T., Klarmann, G.J., Thomas, S.B. & Farrar, W.L., 2008. CD44(+)CD24(-) prostate cells are early cancer progenitor/stem cells that provide a model for patients with poor prognosis. *British Journal of Cancer,* 98(4), pp. 756-765.

Huss, W.J., Gray, D.R., Werdin, E.S., Funkhouser, W.K. & Smith, G.J., 2004. Evidence of pluripotent human prostate stem cells in a human prostate primary xenograft model. *Prostate,* 60(2), pp. 77-90.

Iglesias-Gato, D., Wikstrom, P., Tyanova, S., Lavallee, C., Thysell, E., Carlsson, J., Hagglof, C., Cox, J., Andren, O., Stattin, P., Egevad, L., Widmark, A., Bjartell, A., Collins, C.C., Bergh, A., Geiger, T., Mann, M. & Flores-Morales, A., 2016. The Proteome of Primary Prostate Cancer. *European Urology,* 69(5), pp. 942-952.

Ishibashi, H., Suzuki, T., Suzuki, S., Moriya, T., Kaneko, C., Takizawa, T., Sunamori, M., Handa, M., Kondo, T. & Sasano, H., 2003. Sex steroid hormone receptors in human thymoma. *Journal of Clinical Endocrinology & Metabolism,* 88(5), pp. 2309-2317.

Ishikawa, T. & Kramer, R.H., 2010. Sdc1 negatively modulates carcinoma cell motility and invasion. *Experimental Cell Research,* 316(6), pp. 951-965.

Ismail, A.H.R., Landry, F., Aprikian, A.G. & Chevalier, S., 2002. Androgen ablation promotes neuroendocrine cell differentiation in dog and human prostate. *Prostate,* 51(2).

Iuliano, A., Occhipinti, A., Angelini, C., De Feis, I. & Lio, P., 2016. Cancer Markers Selection Using Network-Based Cox Regression: A Methodological and Computational Practice. *Frontiers in Physiology,* 7.

Jani, J., Cohen, C. & Osunkoya, A.O., 2010. Does the Inverse Immunoexpression Pattern of Nestin and Androgen Receptor Predict Treatment Failure in Patients with Prostatic Adenocarcinoma? *Laboratory Investigation,* 90.

Jenkins, R.B., Qian, J.Q., Lieber, M.M. & Bostwick, D.G., 1997. Detection of c-myc oncogene amplification and chromosomal anomalies in metastatic prostatic carcinoma by fluorescence in situ hybridization. *Cancer Research,* 57(3), pp. 524-531.

Jeong, H.-W., Cui, W., Yang, Y., Lu, J., He, J., Li, A., Song, D., Guo, Y., Liu, B.H. & Chai, L., 2011. SALL4, a Stem Cell Factor, Affects the Side Population by Regulation of the ATP-Binding Cassette Drug Transport Genes. *Plos One,* 6(4).

Jeter, C.R., Liu, B., Liu, X., Chen, X., Liu, C., Calhoun-Davis, T., Repass, J., Zaehres, H., Shen, J.J. & Tang, D.G., 2011. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene,* 30(36).

Jiang, J.F., Jia, P.L., Zhao, Z.M. & Shen, B.R., 2014. Key regulators in prostate cancer identified by co-expression module analysis. *Bmc Genomics,* 15.

Jiang, N., Zhu, S.M., Chen, J., Niu, Y.J. & Zhou, L.Q., 2013. A-Methylacyl-CoA Racemase (AMACR) and Prostate-Cancer Risk: A Meta-Analysis of 4,385 Participants. *Plos One,* 8(10).

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D. & Slack, F.J., 2005. RAS is regulated by the let-7 MicroRNA family. *Cell,* 120(5), pp. 635-647.

Jong, N.N. & McKeage, M.J., 2014. Emerging roles of metal solute carriers in cancer mechanisms and treatment. *Biopharmaceutics & Drug Disposition,* 35(8), pp. 450-462.

Joshi, M.D., Ahmad, R., Yin, L., Raina, D., Rajabi, H., Bubley, G., Kharbanda, S. & Kufe, D., 2009. MUC1 oncoprotein is a druggable target in human prostate cancer cells. *Molecular Cancer Therapeutics,* 8(11), pp. 3056-3065.

Juuti, A., Nordling, S., Lundin, J., Louhimo, J. & Haglund, C., 2005. Syndecan-1 expression - A novel prognostic marker in pancreatic cancer. *Oncology,* 68(2-3), pp. 97-106.

Kang, G.H., Lee, S., Lee, H.J. & Hwang, K.S., 2004. Aberrant CpG island hypermethylation of multiple genes in prostate cancer and prostatic intraepithelial neoplasia. *Journal of Pathology,* 202(2), pp. 233-240.

Kasza, I., Suh, Y., Wollny, D., Clark, R.J., Roopra, A., Colman, R.J., MacDougald, O.A., Shedd, T.A., Nelson, D.W., Yen, M.I., Yen, C.L.E. & Alexander, C.M., 2014. Syndecan-1 Is Required to Maintain Intradermal Fat and Prevent Cold Stress. *Plos Genetics,* 10(8).

Kee, K., Foster, B.A., Merali, S., Kramer, D.L., Hensen, M.L., Diegelman, P., Kisiel, N., Vujcic, S., Mazurchuk, R.V. & Porter, C.W., 2004. Activated polyamine catabolism depletes acetyl-CoA pools and

suppresses prostate tumor growth in TRAMP mice. *Journal of Biological Chemistry,* 279(38), pp. 40076-40083.

Keniry, M. & Parsons, R., 2008. The role of PTEN signaling perturbations in cancer and in targeted therapy. *Oncogene,* 27(41), pp. 5477-5485.

Kim, Y., Jeon, J., Mejia, S., Yao, C.Q., Ignatchenko, V., Nyalwidhe, J.O., Gramolini, A.O., Lance, R.S., Troyer, D.A., Drake, R.R., Boutros, P.C., Semmes, O.J. & Kislinger, T., 2016. Targeted proteomics identifies liquid-biopsy signatures for extracapsular prostate cancer. *Nature Communications,* 7.

King, C.R. & Long, J.P., 2000. Prostate biopsy grading errors: A sampling problem? *International Journal of Cancer,* 90(6), pp. 326-330.

Kleeberger, W., Bova, G.S., Nielsen, M.E., Herawi, M., Chuang, A.-Y., Epstein, J.I. & Berman, D.M., 2007a. Roles for the stem cell-associated intermediate filament nestin in prostate cancer migration and metastasis. *Cancer Research,* 67(19).

Kleeberger, W., Bova, G.S., Nielsen, M.E., Herawi, M., Chuang, A.-Y., Epstein, J.I. & Berman, D.M., 2007b. Roles for the stem cell-associated intermediate filament nestin in prostate cancer migration and metastasis. *Cancer Research,* 67(19), pp. 9199-9206.

Ko, S.B.H., Azuma, S., Yokoyama, Y., Yamamoto, A., Kyokane, K., Niida, S., Ishiguro, H. & Ko, M.S.H., 2013. Inflammation increases cells expressing ZSCAN4 and progenitor cell markers in the adult pancreas. *American Journal of Physiology-Gastrointestinal and Liver Physiology,* 304(12), pp. G1103-G1116.

Kobayashi, D., Kuribayashi, K., Tanaka, M. & Watanabe, N., 2011a. Overexpression of SALL4 in lung cancer and its importance in cell proliferation. *Oncology Reports,* 26(4).

Kobayashi, D., Kuribayshi, K., Tanaka, M. & Watanabe, N., 2011b. SALL4 is essential for cancer cell proliferation and is overexpressed at early clinical stages in breast cancer. *International Journal of Oncology,* 38(4).

Koie, T., Ohyama, C., Hatakeyama, S., Imai, A., Yoneyama, T., Hashimoto, Y., Tobisawa, Y., Hosogoe, S., Yamamoto, H., Kitayama, M. & Hirota, K., 2016. Significance of preoperative butyrylcholinesterase as an independent predictor of biochemical recurrence-free survival in patients with prostate cancer treated with radical prostatectomy. *International Journal of Clinical Oncology,* 21(2), pp. 379-383.

Kong, D.J., Banerjee, S., Ahmad, A., Li, Y.W., Wang, Z.W., Sethi, S. & Sarkar, F.H., 2010. Epithelial to Mesenchymal Transition Is Mechanistically Linked with Stem Cell Signatures in Prostate Cancer Cells. *Plos One,* 5(8), p. 14.

Kong, Y., Peng, Y.Q., Liu, Y., Xin, H., Zhan, X. & Tan, W.F., 2015. Twist1 and Snail Link Hedgehog Signaling to Tumor-Initiating Cell-Like Properties and Acquired Chemoresistance Independently of ABC Transporters. *Stem Cells,* 33(4), pp. 1063-1074.

Koo, K.M., Wee, E.J.H., Mainwaring, P.N. & Trau, M., 2016. A simple, rapid, low-cost technique for naked-eye detection of urine-isolated TMPRSS2: ERG gene fusion RNA. *Scientific Reports,* 6.

Kregel, S., Kiriluk, K.J., Rosen, A.M., Cai, Y., Reyes, E.E., Otto, K.B., Tom, W., Paner, G.P., Szmulewitz, R.Z. & Vander Griend, D.J., 2013. Sox2 is an androgen receptor-repressed gene that promotes castration-resistant prostate cancer. *PLoS One.* United States, p. e53701.

Krijnen, J.L.M., Janssen, P.J.A., Dewinter, J.A.R., Vankrimpen, H., Schroder, F.H. & Vanderkwast, T.H., 1993. DO NEUROENDOCRINE CELLS IN HUMAN PROSTATE-CANCER EXPRESS ANDROGEN RECEPTOR. *Histochemistry,* 100(5), pp. 393-398.

Kuczyk, M.A., Serth, J., Bokemeyer, C., Machtens, S., Minssen, A., Bathke, W., Hartmann, J. & Jonas, U., 1998. The prognostic value of p53 for long-term and recurrence-free survival following radical prostatectomy. *European Journal of Cancer,* 34(5), pp. 679-686.

Kumar, B., de Silva, M., Venter, D.J. & Armes, J.E., 2004. Tissue microarrays: a practical guide. *Pathology,* 36(4), pp. 295-300.

Kundu, S.D., Roehl, K.A., Eggener, S.E., Antenor, J.A.V., Han, M. & Catalona, W.J., 2004. Potency, continence and complications in 3,477 consecutive radical retropubic prostatectomies. *Journal of Urology,* 172(6), pp. 2227-2231.

Kurita, T., Medina, R.T., Mills, A.A. & Cunha, G.R., 2004. Role of p63 and basal cells in the prostate. *Development,* 131(20), pp. 4955-4964.

Kurita, T., Wang, Y.Z., Donjacour, A.A., Zhao, C., Lydon, J.P., O'Malley, B.W., Isaacs, J.T., Dahiya, R. & Cunha, G.R., 2001. Paracrine regulation of apoptosis by steroid hormones in the male and female reproductive system. *Cell Death and Differentiation,* 8(2), pp. 192-200.

Kyprianou, N. & Isaacs, J.T., 1988. ACTIVATION OF PROGRAMMED CELL-DEATH IN THE RAT VENTRAL PROSTATE AFTER CASTRATION. *Endocrinology,* 122(2), pp. 552-562.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceresc ortes, J., Minden, M., Paterson, B., Caligiuri, M.A. & Dick, J.E., 1994. A cell initiating human acute myeloid-leukemia after transplantation into SCID mice. *Nature,* 367(6464), pp. 645-648.

Lapointe, J., Li, C., Higgins, J.P., van de Rijn, M., Bair, E., Montgomery, K., Ferrari, M., Egevad, L., Rayford, W., Bergerheim, U., Ekman, P., DeMarzo, A.M., Tibshirani, R., Botstein, D., Brown, P.O., Brooks, J.D. & Pollack, J.R., 2004. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America,* 101(3), pp. 811-816.

Lawson, D.A., Zong, Y., Memarzadeh, S., Xin, L., Huang, J. & Witte, O.N., 2010. Basal epithelial stem cells are efficient targets for prostate cancer initiation. *Proceedings of the National Academy of Sciences of the United States of America,* 107(6).

Leclerc, B.G., Charlebois, R., Chouinard, G., Allard, B., Pommey, S., Saad, F. & Stagg, J., 2016. CD73 Expression Is an Independent Prognostic Factor in Prostate Cancer. *Clinical Cancer Research,* 22(1), pp. 158-166.

Ledezma, R., Cifuentes, F., Gallegos, I., Fulla, J., Ossandon, E., Castellon, E.A. & Contreras, H.R., 2011. Altered expression patterns of syndecan-1 and -2 predict biochemical recurrence in prostate cancer. *Asian Journal of Andrology,* 13(3).

Lee, B.H., Kibel, A.S., Ciezki, J.P., Klein, E.A., Reddy, C.A., Yu, C.H., Kattan, M.W. & Stephenson, A.J., 2015a. Are Biochemical Recurrence Outcomes Similar After Radical Prostatectomy and Radiation Therapy? Analysis of Prostate Cancer-Specific Mortality by Nomogram-predicted Risks of Biochemical Recurrence. *European Urology,* 67(2), pp. 204-209.

Lee, K. & Gollahon, L.S., 2014. Zscan4 interacts directly with human Rap1 in cancer cells regardless of telomerase status. *Cancer Biology & Therapy,* 15(8), pp. 1094-1105.

Lee, K. & Gollahon, L.S., 2015. ZSCAN4 and TRF1: A functionally indirect interaction in cancer cells independent of telomerase activity. *Biochemical and Biophysical Research Communications,* 466(4), pp. 644-649.

Lee, S.H., Choi, E.J., Kim, M.S., Park, J.W., Lee, Y.S., Kim, S.Y. & Kang, C.S., 2014. Prognostic significance of syndecan-1 expression in squamous cell carcinoma of the tonsil. *International Journal of Clinical Oncology,* 19(2), pp. 247-253.

Lee, S.H., Johnson, D.T., Luong, R., Yu, E.J., Cunha, G.R., Nusse, R. & Sun, Z.J., 2015b. Wnt/beta-Catenin-Responsive Cells in Prostatic Development and Regeneration. *Stem Cells,* 33(11), pp. 3356-3367.

Lee, Y.S. & Dutta, A., 2007. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes & Development,* 21(9), pp. 1025-1030.

Leibovici, D., Spiess, P.E., Agarwal, P.K., Tu, S.M., Pettaway, C.A., Hitzhusen, K., Millikan, R.E. & Pisters, L.L., 2007. Prostate cancer progression in the presence of undetectable or low serum prostate-specific antigen level. *Cancer,* 109(2), pp. 198-204.

Leinonen, K.A., Tolonen, T.T., Bracken, H., Stenman, U.H., Tammela, T.L.J., Saramaki, O.R. & Visakorpi, T., 2010. Association of SPINK1 Expression and TMPRSS2:ERG Fusion with Prognosis in Endocrine-Treated Prostate Cancer. *Clinical Cancer Research,* 16(10), pp. 2845-2851.

Leong, K.G., Wang, B.E., Johnson, L. & Gao, W.Q., 2008. Generation of a prostate from a single adult stem cell. *Nature,* 456(7223), pp. 804-808.

Levine, A.J., 1997. p53, the cellular gatekeeper for growth and division. *Cell,* 88(3), pp. 323-331.

Leyten, G., Hessels, D., Jannink, S.A., Smit, F.P., de Jong, H., Cornel, E.B., de Reijke, T.M., Vergunst, H., Kil, P., Knipscheer, B.C., van Oort, I.M., Mulders, P.F.A., Hulsbergen-van de Kaa, C.A. & Schalken, J.A., 2014. Prospective Multicentre Evaluation of PCA3 and TMPRSS2-ERG Gene Fusions as Diagnostic and Prognostic Urinary Biomarkers for Prostate Cancer. *European Urology,* 65(3), pp. 534-542.

Li, C., Ao, J., Fu, J., Lee, D.F., Xu, J., Lonard, D. & O'Malley, B.W., 2011. Tumor-suppressor role for the SPOP ubiquitin ligase in signal-dependent proteolysis of the oncogenic co-activator SRC-3/AIB1. *Oncogene,* 30(42), pp. 4350-4364.

Li, T., Su, Y., Mei, Y.P., Leng, Q.X., Leng, B.J., Liu, Z.Q., Stass, S.A. & Jiang, F., 2010. ALDH1A1 is a marker for malignant prostate stem cells and predictor of prostate cancer patients' outcome. *Laboratory Investigation,* 90(2), pp. 234-244.

Lilja, H., Ulmert, D. & Vickers, A.J., 2008. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. *Nature Reviews Cancer,* 8(4).

Lin, X.H., Tascilar, M., Lee, W.H., Vles, W.J., Lee, B.H., Veeraswamy, R., Asgari, K., Freije, D., van Rees, B., Gage, W.R., Bova, G.S., Isaacs, W.B., Brooks, J.D., DeWeese, T.L., De Marzo, A.M. & Nelson, W.G., 2001. GSTP1 CpG island hypermethylation is responsible for the absence of GSTP1 expression in human prostate cancer cells. *American Journal of Pathology,* 159(5), pp. 1815-1826.

Liu, A.J., Wei, L.X., Gardner, W.A., Deng, C.X. & Man, Y.G., 2009. Correlated alterations in prostate basal cell layer and basement membrane. *International Journal of Biological Sciences,* 5(3), pp. 276-285.

Liu, B.Y., McDermott, S.P., Khwaja, S.S. & Alexander, C.M., 2004. The transforming activity of Wnt effectors correlates with their ability to induce the accumulation of mammary progenitor cells. *Proceedings of the National Academy of Sciences of the United States of America,* 101(12), pp. 4158-4163.

Liu, C., Kelnar, K., Liu, B.G., Chen, X., Calhoun-Davis, T., Li, H.W., Patrawala, L., Yan, H., Jeter, C., Honorio, S., Wiggins, J.F., Bader, A.G., Fagin, R., Brown, D. & Tang, D.A.G., 2011a. The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Nature Medicine,* 17(2), pp. 211-U105.

Liu, C., Kelnar, K., Vlassov, A.V., Brown, D., Wang, J.C. & Tang, D.G., 2012. Distinct microRNA Expression Profiles in Prostate Cancer Stem/Progenitor Cells and Tumor-Suppressive Functions of let-7. *Cancer Research,* 72(13), pp. 3393-3404.

Liu, Y.-L., Yang, Y.-M., Xu, H. & Dong, X.-S., 2011b. Aberrant Expression of USP22 is Associated With Liver Metastasis and Poor Prognosis of Colorectal Cancer. *Journal of Surgical Oncology,* 103(3).

Lotan, T.L., Gumuskaya, B., Rahimi, H., Hicks, J.L., Iwata, T., Robinson, B.D., Epstein, J.I. & De Marzo, A.M., 2013. Cytoplasmic PTEN protein loss distinguishes intraductal carcinoma of the prostate from high-grade prostatic intraepithelial neoplasia. *Modern Pathology,* 26(4), pp. 587-603.

Lukacs, R.U., Goldstein, A.S., Lawson, D.A., Cheng, D. & Witte, O.N., 2010a. Isolation, cultivation and characterization of adult murine prostate stem cells. *Nature Protocols,* 5(4), pp. 702-713.

Lukacs, R.U., Memarzadeh, S., Wu, H. & Witte, O.N., 2010b. Bmi-1 Is a Crucial Regulator of Prostate Stem Cell Self-Renewal and Malignant Transformation. *Cell Stem Cell,* 7(6).

Luo, J., Zha, S., Gage, W.R., Dunn, T.A., Hicks, J.L., Bennett, C.J., Ewing, C.N., Platz, E.A., Ferdinandusse, S., Wanders, R.J., Trent, J.M., Isaacs, W.B. & De Marzo, A.M., 2002. alpha-Methylacyl-CoA racemase: A new molecular marker for prostate cancer. *Cancer Research,* 62(8), pp. 2220-2226.

Ma, Y.Y., Liang, D.M., Liu, J., Axcrona, K., Kvalheim, G., Stokke, T., Nesland, J.M. & Suo, Z.H., 2011. Prostate Cancer Cell Lines under Hypoxia Exhibit Greater Stem-Like Properties. *Plos One,* 6(12), p. 13.

Machtens, S., Serth, J., Bokemeyer, C., Bathke, W., Minssen, A., Kollmannsberger, C., Hartmann, J., Knuchel, R., Kondo, M., Jonas, U. & Kuczyk, M., 2001. Expression of the p53 and maspin protein in primary prostate cancer: Correlation with clinical features. *International Journal of Cancer,* 95(5), pp. 337-342.

Magi-Galluzzi, C., Maddala, T., Falzarano, S.M., Cherbavaz, D.B., Zhang, N., Knezevic, D., Febbo, P.G., Lee, M., Lawrence, H.J. & Klein, E.A., 2016. Gene expression in normal-appearing tissue adjacent to prostate cancers are predictive of clinical outcome: evidence for a biologically meaningful field effect. *Oncotarget,* [Epub ahead of print].

Maiga, R.I., Lemieux, J., Roy, A., Simard, C. & Neron, S., 2014. Flow Cytometry Assessment of In Vitro Generated CD138(+) Human Plasma Cells. *Biomed Research International*.

Majumder, P.K., Grisanzio, C., O'Connell, F., Barry, M., Brito, J.M., Xu, Q., Guney, I., Berger, R., Herman, P., Bikoff, R., Fedele, G., Baek, W.K., Wang, S., Ellwood-Yen, K., Wu, H., Sawyers, C.L., Signoretti, S., Hahn, W.C., Loda, M. & Sellers, W.R., 2008. A prostatic intraepithelial neoplasia-dependent p27(Kip1) checkpoint induces senescence and inhibits cell proliferation and cancer progression. *Cancer Cell,* 14(2), pp. 146-155.

Mani, R.S., Tomlins, S.A., Callahan, K., Ghosh, A., Nyati, M.K., Varambally, S., Palanisamy, N. & Chinnaiyan, A.M., 2009. Induced Chromosomal Proximity and Gene Fusions in Prostate Cancer. *Science,* 326(5957), pp. 1230-1230.

Markert, E.K., Mizuno, H., Vazquez, A. & Levine, A.J., 2011. Molecular classification of prostate cancer using curated expression signatures. *Proceedings of the National Academy of Sciences of the United States of America,* 108(52), pp. 21276-21281.

Massie, C.E., Lynch, A., Ramos-Montoya, A., Boren, J., Stark, R., Fazli, L., Warren, A., Scott, H., Madhu, B., Sharma, N., Bon, H., Zecchini, V., Smith, D.M., DeNicola, G.M., Mathews, N., Osborne, M., Hadfield, J., MacArthur, S., Adryan, B., Lyons, S.K., Brindle, K.M., Griffiths, J., Gleave, M.E., Rennie, P.S., Neal, D.E. & Mills, I.G., 2011. The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis. *Embo Journal,* 30(13), pp. 2719-2733.

Mathew, G., Timm, E.A., Sotomayor, P., Godoy, A., Montecinos, V.P., Smith, G.J. & Huss, W.J., 2009. ABCG2-mediated dyecycle violet efflux defined side population in benign and malignant prostate. *Cell Cycle,* 8(7), pp. 1053-1061.

Mathews, L.A., Hurt, E.M., Zhang, X. & Farrar, W.L., 2010. Epigenetic regulation of CpG promoter methylation in invasive prostate cancer cells. *Molecular Cancer,* 9.

Mathieu, J., Zhang, Z., Zhou, W.Y., Wang, A.J., Heddleston, J.M., Pinna, C.M.A., Hubaud, A., Stadler, B., Choi, M., Bar, M., Tewari, M., Liu, A., Vessella, R., Rostomily, R., Born, D., Horwitz, M., Ware, C., Blau, C.A., Cleary, M.A., Rich, J.N. & Ruohola-Baker, H., 2011. HIF Induces Human Embryonic Stem Cell Markers in Cancer Cells. *Cancer Research,* 71(13), pp. 4640-4652.

Matilainen, H., Yu, X.W., Tang, C.W., Berridge, M.V. & McConnell, M.J., 2012. Sphere formation reverses the metastatic and cancer stem cell phenotype of the murine mammary tumour 4T1, independently of the putative cancer stem cell marker Sca-1. *Cancer Letters,* 323(1), pp. 20-28.

Mavropoulos, A., Simopoulou, T., Varna, A., Liaskos, C., Katsiari, C.G., Bogdanos, D.P. & Sakkas, L.I., 2016. Breg Cells Are Numerically Decreased and Functionally Impaired in Patients With Systemic Sclerosis. *Arthritis & Rheumatology,* 68(2), pp. 494-504.

McDonnell, T.J., Navone, N.M., Troncoso, P., Pisters, L.L., Conti, C., vonEschenbach, A.C., Brisbay, S. & Logothetis, C.J., 1997. Expression of bcl-2 oncoprotein and p53 protein accumulation in bone marrow metastases of androgen independent prostate cancer. *Journal of Urology,* 157(2), pp. 569-574.

McDonnell, T.J., Troncoso, P., Brisbay, S.M., Logothetis, C., Chung, L.W.K., Hsieh, J.T., Tu, S.M. & Campbell, M.L., 1992. EXPRESSION OF THE PROTOONCOGENE BCL-2 IN THE PROSTATE AND ITS ASSOCIATION WITH EMERGENCE OF ANDROGEN-INDEPENDENT PROSTATE-CANCER. *Cancer Research,* 52(24), pp. 6940-6944.

McNeal, J.E., 1981. THE ZONAL ANATOMY OF THE PROSTATE. *Prostate,* 2(1), pp. 35-49.

McNeal, J.E., 1988. NORMAL HISTOLOGY OF THE PROSTATE. *American Journal of Surgical Pathology,* 12(8), pp. 619-633.

McNeal, J.E., Haillot, O. & Yemoto, C., 1995. CELL-PROLIFERATION IN DYSPLASIA OF THE PROSTATE - ANALYSIS BY PCNA IMMUNOSTAINING. *Prostate,* 27(5), pp. 258-268.

McNeal, J.E., Redwine, E.A., Freiha, F.S. & Stamey, T.A., 1988. ZONAL DISTRIBUTION OF PROSTATIC ADENOCARCINOMA - CORRELATION WITH HISTOLOGIC PATTERN AND DIRECTION OF SPREAD. *American Journal of Surgical Pathology,* 12(12), pp. 897-906.

Medrano, A., Fernandez-Novell, J.M., Ramio, L., Alvarez, J., Goldberg, E., Rivera, M.M., Guinovart, J.J., Rigau, T. & Rodriguez-Gil, J.E., 2006. Utilization of citrate and lactate through a lactate dehydrogenase and ATP-regulated pathway in boar spermatozoa. *Molecular Reproduction and Development,* 73(3), pp. 369-378.

Melia, J., Moseley, R., Ball, R.Y., Griffiths, D.F.R., Grigor, K., Harnden, P., Jarmulowicz, M., McWilliam, L.J., Montironi, R., Waller, M., Moss, S. & Parkinson, M.C., 2006. A UK-based investigation of inter- and intra-observer reproducibility of Gleason grading of prostatic biopsies. *Histopathology,* 48(6), pp. 644-654.

Melling, N., Harutyunyan, L., Hube-Magg, C., Kluth, M., Simon, R., Lebok, P., Minner, S., Tsourlakis, M.C., Koop, C., Graefen, M., Adam, M., Haese, A., Wittmer, C., Steurer, S., Izbicki, J., Sauter, G., Wilczak, W., Schlomm, T. & Krech, T., 2015. High-Level HOOK3 Expression Is an Independent Predictor of Poor Prognosis Associated with Genomic Instability in Prostate Cancer. *Plos One,* 10(7).

Mennerich, D., Vogel, A., Klaman, I., Dahl, E., Lichtner, R.B., Rosenthal, A., Pohlenz, H.D., Thierauch, K.H. & Sommer, A., 2004. Shift of syndecan-1 expression from epithelial to stromal cells during progression of solid tumours. *European Journal of Cancer,* 40(9), pp. 1373-1382.

Mercader, M., Bodner, B.K., Moser, M.T., Kwon, P.S., Park, E.S.Y., Manecke, R.G., Ellis, T.M., Wojcik, E.M., Yang, D., Flanigan, R.C., Waters, W.B., Kast, W.M. & Kwon, E.D., 2001. T cell infiltration of the prostate induced by androgen withdrawal in patients with prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America,* 98(25), pp. 14565-14570.

Merola, R., Tomao, L., Antenucci, A., Sperduti, I., Sentinelli, S., Masi, S., Mandoj, C., Orlandi, G., Papalia, R., Guaglianone, S., Costantini, M., Cusumano, G., Cigliana, G., Ascenzi, P., Gallucci, M. & Conti, L., 2015. PCA3 in prostate cancer and tumor aggressiveness detection on 407 high-risk patients: a National Cancer Institute experience. *Journal of Experimental & Clinical Cancer Research,* 34.

Merseburger, A.S., Kuczyk, M.A., Serth, J., Bokemeyer, C., Young, D.Y., Sun, L., Connelly, R.R., McLeod, D.G., Mostofi, F.K., Srivastava, S.K., Stenzl, A., Moul, J.W. & Sesterhenn, I.A., 2003. Limitations of tissue microarrays in the evaluation of focal alterations of bcl-2 and p53 in whole mount derived prostate tissues. *Oncology Reports,* 10(1), pp. 223-228.

Meyers, F.J., Gumerlock, P.H., Chi, S.G., Borchers, H., Deitch, A.D. & White, R.W.D., 1998. Very frequent p53 mutations in metastatic prostate carcinoma and in matched primary tumors. *Cancer,* 83(12), pp. 2534-2539.

Miki, J., Furusato, B., Li, H.Z., Gu, Y.P., Takahashi, H., Egawa, S., Sesterhenn, I.A., McLeod, D.G., Srivastava, S. & Rhim, J.S., 2007. Identification of putative stem cell markers, CD133 and CXCR4, in hTERT-immortalized primary nonmalignant and malignant tumor-derived human prostate epithelial cell lines and in prostate cancer specimens. *Cancer Research,* 67(7), pp. 3153-3161.

Mucci, N.R., Akdas, G., Manely, S. & Rubin, M.A., 2000. Neuroendocrine expression in metastatic prostate cancer: Evaluation of high throughput tissue microarrays to detect heterogeneous protein expression. *Human Pathology,* 31(4), pp. 406-414.

Mukouyama, H., Niimura, K. & Ogawa, Y., 2001. Histologic upgrading of prostate cancer occurs frequently over a short period of time: Single hospital experiences of radical prostatectomy. *International Journal of Urology,* 8(8), pp. S45-S49.

Murohashi, M., Hinohara, K., Kuroda, M., Isagawa, T., Tsuji, S., Kobayashi, S., Umezawa, K., Tojo, A., Aburatani, H. & Gotoh, N., 2010. Gene set enrichment analysis provides insight into novel signalling pathways in breast cancer stem cells. *British Journal of Cancer,* 102(1).

Nadendla, S.K., Hazan, A., Ward, M., Harper, L.J., Moutasim, K., Bianchi, L.S., Naase, M., Ghali, L., Thomas, G.J., Prowse, D.M., Philpott, M.P. & Neill, G.W., 2011. GLI1 Confers Profound Phenotypic Changes upon LNCaP Prostate Cancer Cells That Include the Acquisition of a Hormone Independent State. *Plos One,* 6(5).

Nakai-Futatsugi, Y. & Niwa, H., 2016. Zscan4 Is Activated after Telomere Shortening in Mouse Embryonic Stem Cells. *Stem Cell Reports,* 6(4), pp. 483-495.

Navone, N.M., Troncoso, P., Pisters, L.L., Goodrow, T.L., Palmer, J.L., Nichols, W.W., Voneschenbach, A.C. & Conti, C.J., 1993. P53 PROTEIN ACCUMULATION AND GENE MUTATION IN THE PROGRESSION OF HUMAN PROSTATE CARCINOMA. *Journal of the National Cancer Institute,* 85(20), pp. 1657-1669.

Nguyen, T.L., Grizzle, W.E., Zhang, K., Hameed, O., Siegal, G.P. & Wei, S., 2013. Syndecan-1 Overexpression Is Associated With Nonluminal Subtypes and Poor Prognosis in Advanced Breast Cancer. *American Journal of Clinical Pathology,* 140(4), pp. 468-474.

Ni, J., Cozzi, P.J., Hao, J.L., Beretov, J., Chang, L., Duan, W., Shigdar, S., Delprado, W.J., Graham, P.H., Bucci, J., Kearsley, J.H. & Li, Y., 2014. CD44 variant 6 is associated with prostate cancer metastasis and chemo-/radioresistance. *Prostate,* 74(6), pp. 602-617.

Nickel, J.C., Roehrborn, C.G., O'Leary, M.P., Bostwick, D.G., Somerville, M.C. & Rittmaster, R.S., 2008. The Relationship between Prostate Inflammation and Lower Urinary Tract Symptoms: Examination of Baseline Data from the REDUCE Trial. *European Urology,* 54(6), pp. 1379-1384.

Nikonova, A.S., Astsaturov, I., Serebriiskii, I.G., Dunbrack, R.L. & Golemis, E.A., 2013. Aurora A kinase (AURKA) in normal and pathological cell division. *Cellular and Molecular Life Sciences,* 70(4), pp. 661-687.

Oh, S.-Y., Kang, H.J., Kim, Y.S., Kim, H. & Lim, Y.C., 2013. CD44-negative cells in head and neck squamous carcinoma also have stem-cell like traits. *European journal of cancer (Oxford, England : 1990),* 49(1).

Ou, J., Vijayasarathy, C., Ziccardi, L., Chen, S., Zeng, Y., Marangoni, D., Pope, J.G., Bush, R.A., Wu, Z., Li, W. & Sieving, P.A., 2015. Synaptic pathology and therapeutic repair in adult retinoschisis mouse by AAV-RS1 transfer. *Journal of Clinical Investigation,* 125(7), pp. 2891-2903.

Ousset, M., Van Keymeulen, A., Bouvencourt, G., Sharma, N., Achouri, Y., Simons, B.D. & Blanpain, C., 2012. Multipotent and unipotent progenitors contribute to prostate postnatal development. *Nature Cell Biology,* 14(11), pp. 1130-+.

Ouyang, X., DeWeese, T.L., Nelson, W.G. & Abate-Shen, C., 2005. Loss-of-function of Nkx3.1 promotes increased oxidative damage in prostate carcinogenesis. *Cancer Research,* 65(15), pp. 6773-6779.

Palmer, N.P., Schmid, P.R., Berger, B. & Kohane, I.S., 2012. A gene expression profile of stem cell pluripotentiality and differentiation is conserved across diverse solid and hematopoietic cancers. *Genome Biology,* 13(8).

Pan, X.Y., Zhang, X.M., Gong, J., Tan, J.Y., Yin, X.X., Tang, Q.D., Shu, K.P., Shen, P.F., Zeng, H. & Chen, N., 2016. The Expression Profile and Prognostic Value of SPINK1 in Initially Diagnosed Bone Metastatic Prostate Cancer. *Prostate,* 76(9), pp. 823-833.

Paner, G.P., Luthringer, D.J. & Amin, M.B., 2008. Best practice in diagnostic immunohistochemistry - Prostate carcinoma and its mimics in needle core biopsies. *Archives of Pathology & Laboratory Medicine,* 132(9), pp. 1388-1396.

Pasqualon, T., Lue, H.Q., Groening, S., Pruessmeyer, J., Jahr, H., Denecke, B., Bernhagen, J. & Ludwig, A., 2016. Cell surface syndecan-1 contributes to binding and function of macrophage migration inhibitory factor (MIF) on epithelial tumor cells. *Biochimica Et Biophysica Acta-Molecular Cell Research,* 1863(4), pp. 717-726.

Pastrana, E., Silva-Vargas, V. & Doetsch, F., 2011. Eyes Wide Open: A Critical Review of Sphere-Formation as an Assay for Stem Cells. *Cell Stem Cell,* 8(5), pp. 486-498.

Patrawala, L., Calhoun, T., Schneider-Broussard, R., Li, H., Bhatia, B., Tang, S., Reilly, J.G., Chandra, D., Zhou, J., Claypool, K., Coghlan, L. & Tang, D.G., 2006. Highly purified CD44(+) prostate cancer cells from xenograft human tumors are enriched in tumorigenic and metastatic progenitor cells. *Oncogene,* 25(12).

Patrawala, L., Calhoun, T., Schneider-Broussard, R., Zhou, J.J., Claypool, K. & Tang, D.G., 2005. Side population is enriched in tumorigenic, stem-like cancer cells, whereas ABCG2(+) and ABCG2(-) cancer cells are similarly tumorigenic. *Cancer Research,* 65(14), pp. 6207-6219.

Patrawala, L., Calhoun-Davis, T., Schneider-Broussard, R. & Tang, D.G., 2007. Hierarchical organization of prostate cancer cells in xenograft tumors: The CD44(+)alpha 2 beta 1(+) cell population is enriched in tumor-initiating cells. *Cancer Research,* 67(14), pp. 6796-6805.

Penney, K.L., Stampfer, M.J., Jahn, J.L., Sinnott, J.A., Flavin, R., Rider, J.R., Finn, S., Giovannucci, E., Sesso, H.D., Loda, M., Mucci, L.A. & Fiorentino, M., 2013. Gleason Grade Progression Is Uncommon. *Cancer Research,* 73(16), pp. 5163-5168.

Petrovics, G., Liu, A.J., Shaheduzzaman, S., Furasato, B., Sun, C., Chen, Y.M., Nau, M., Ravindranath, L., Chen, Y.D., Dobi, A., Srikantan, V., Sesterhenn, I.A., McLeod, D.G., Vahey, M., Moul, J.W. & Srivastava, S., 2005. Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. *Oncogene,* 24(23), pp. 3847-3852.

Pfeiffer, M.J. & Schalken, J.A., 2010. Stem Cell Characteristics in Prostate Cancer Cell Lines. *European Urology,* 57(2), pp. 246-254.

Pfeiffer, M.J., Smit, F.P., Sedelaar, J.P.M. & Schalken, J.A., 2011. Steroidogenic Enzymes and Stem Cell Markers Are Upregulated during Androgen Deprivation in Prostate Cancer. *Molecular Medicine,* 17(7-8), pp. 657-664.

Pinder, S.E., Brown, J.P., Gillett, C., Purdie, C.A., Speirs, V., Thompson, A.M., Shaaban, A.M. & Translational Subgroup of the, N.B.C.S.G., 2013. The manufacture and assessment of tissue microarrays: suggestions and criteria for analysis, with breast cancer as an example. *Journal of clinical pathology,* 66(3), pp. 169-77.

Poblete, C.E., Fulla, J., Gallardo, M., Munoz, V., Castellon, E.A., Gallegos, I. & Contreras, H.R., 2014. Increased SNAIL expression and low syndecan levels are associated with high Gleason grade in prostate cancer. *International Journal of Oncology,* 44(3), pp. 647-654.

Ponta, H., Sherman, L. & Herrlich, P.A., 2003. CD44: From adhesion molecules to signalling regulators. *Nature Reviews Molecular Cell Biology,* 4(1), pp. 33-45.

Pound, C.R., Partin, A.W., Eisenberger, M.A., Chan, D.W., Pearson, J.D. & Walsh, P.C., 1999. Natural history of progression after PSA elevation following radical prostatectomy. *Jama-Journal of the American Medical Association,* 281(17), pp. 1591-1597.

Prescott, J., Jariwala, U., Jia, L., Cogan, J.P., Barski, A., Pregizer, S., Shen, H.C., Arasheben, A., Neilson, J.J., Frenkel, B. & Coetzee, G.A., 2007. Androgen receptor-mediated repression of novel target genes. *Prostate,* 67(13), pp. 1371-1383.

Purushothaman, A., Uyama, T., Kobayashi, F., Yamada, S., Sugahara, K., Rapraeger, A.C. & Sanderson, R.D., 2010. Heparanase-enhanced shedding of syndecan-1 by myeloma cells promotes endothelial invasion and angiogenesis. *Blood,* 115(12), pp. 2449-2457.

Qian, J.Q., Hirasawa, K., Bostwick, D.G., Bergstralh, E.J., Slezak, J.M., Anderl, K.L., Borell, T.J., Lieber, M.M. & Jenkins, R.B., 2002. Loss of p53 and c-myc overrepresentation in stage T2-3N1-3M0 prostate cancer are potential markers for cancer progression. *Modern Pathology,* 15(1), pp. 35-44.

Qin, J.C., Liu, X., Laffin, B., Chen, X., Choy, G., Jeter, C.R., Calhoun-Davis, T., Li, H.W., Palapattu, G.S., Pang, S., Lin, K., Huang, J.T., Ivanov, I., Li, W., Suraneni, M.V. & Tang, D.G., 2012. The PSA(-/Io) Prostate Cancer Cell Population Harbors Self-Renewing Long-Term Tumor-Propagating Cells that Resist Castration. *Cell Stem Cell,* 10(5), pp. 556-569.

Quiring, R., Wittbrodt, B., Henrich, T., Ramialison, M., Burgtorf, C., Lehrach, H. & Wittbrodt, J., 2004. Large-scale expression screening by automated whole-mount in situ hybridization. *Mechanisms of Development,* 121(7-8), pp. 971-976.

Rajabi, H., Ahmad, R., Jin, C., Joshi, M.D., Guha, M., Alam, M., Kharbanda, S. & Kufe, D., 2012. MUC1-C oncoprotein confers androgen-independent growth of human prostate cancer cells. *Prostate,* 72(15), pp. 1659-1668.

Rajan, P., Stockley, J., Sudbery, I.M., Fleming, J.T., Hedley, A., Kalna, G., Sims, D., Ponting, C.P., Heger, A., Robson, C.N., McMenemin, R.M., Pedley, I.D. & Leung, H.Y., 2014. Identification of a candidate prognostic gene signature by transcriptome analysis of matched pre-and post-treatment prostatic biopsies from patients with advanced prostate cancer. *Bmc Cancer,* 14.

Rajasekhar, V.K., Studer, L., Gerald, W., Socci, N.D. & Scher, H.I., 2011. Tumour-initiating stem-like cells in human prostate cancer exhibit increased NF-kappa B signalling. *Nature Communications,* 2, p. 13.

Ramani, V.C. & Sanderson, R.D., 2014. Chemotherapy stimulates syndecan-1 shedding: A potentially negative effect of treatment that may promote tumor relapse. *Matrix Biology,* 35, pp. 215-222.

Ramos-Vara, J.A., 2005. Technical aspects of immunohistochemistry. *Veterinary Pathology,* 42(4), pp. 405-426.

Reynier, F., Petit, F., Paye, M., Turrel-Davin, F., Imbert, P.E., Hot, A., Mougin, B. & Miossec, P., 2011. Importance of Correlation between Gene Expression Levels: Application to the Type I Interferon Signature in Rheumatoid Arthritis. *Plos One,* 6(10).

Reynolds, B.A. & Weiss, S., 1992. Generation of neurons and astrocytes from isolated cells of the adult mammalian central-nervous-system. *Science,* 255(5052), pp. 1707-1710.

Rhodes, D.R., Barrette, T.R., Rubin, M.A., Ghosh, D. & Chinnaiyan, A.M., 2002. Meta-analysis of microarrays: Interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Research,* 62(15), pp. 4427-4433.

Richardson, G.D., Robson, C.N., Lang, S.H., Neal, D.E., Maitland, N.J. & Collins, A.T., 2004. CD133, a novel marker for human prostatic epithelial stem cells. *Journal of Cell Science,* 117(16), pp. 3539-3545.

Roach, M., Hanks, G., Thames, H., Schellhammer, P., Shipley, W.U., Sokol, G.H. & Sandler, H., 2006. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: Recommendations of the RTOG-ASTRO Phoenix Consensus Conference. *International Journal of Radiation Oncology Biology Physics,* 65(4), pp. 965-974.

Robinson, D., Van Allen, E.M., Wu, Y.-M., Schultz, N., Lonigro, R.J., Mosquera, J.-M., Montgomery, B., Taplin, M.-E., Pritchard, C.C., Attard, G., Beltran, H., Abida, W., Bradley, R.K., Vinson, J., Cao, X., Vats, P., Kunju, L.P., Hussain, M., Feng, F.Y., Tomlins, S.A., Cooney, K.A., Smith, D.C., Brennan, C., Siddiqui, J., Mehra, R., Chen, Y., Rathkopf, D.E., Morris, M.J., Solomon, S.B., Durack, J.C., Reuter, V.E., Gopalan, A., Gao, J., Loda, M., Lis, R.T., Bowden, M., Balk, S.P., Gaviola, G., Sougnez, C., Gupta, M., Yu, E.Y., Mostaghel, E.A., Cheng, H.H., Mulcahy, H., True, L.D., Plymate, S.R., Dvinge, H., Ferraldeschi, R., Flohr, P., Miranda, S., Zafeiriou, Z., Tunariu, N., Mateo, J., Perez-Lopez, R., Demichelis, F., Robinson, B.D., Schiffman, M., Nanus, D.M., Tagawa, S.T., Sigaras, A., Eng, K.W., Elemento, O., Sboner, A., Heath, E.I., Scher, H.I., Pienta, K.J., Kantoff, P., de Bono, J.S., Rubin, M.A., Nelson, P.S., Garraway, L.A., Sawyers, C.L. & Chinnaiyan, A.M., 2015. Integrative Clinical Genomics of Advanced Prostate Cancer. *Cell,* 161(5), pp. 1215-1228.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. & Mesirov, J.P., 2011. Integrative genomics viewer. *Nature Biotechnology,* 29(1), pp. 24-26.

Roudier, M.P., Winters, B.R., Coleman, I., Lam, H.M., Zhang, X.T., Coleman, R., Chery, L., True, L.D., Higano, C.S., Montgomery, B., Lange, P.H., Snyder, L.A., Srivastava, S., Corey, E., Vessella, R.L., Nelson, P.S., Uren, A. & Morrissey, C., 2016. Characterizing the Molecular Features of ERG-Positive Tumors in Primary and Castration Resistant Prostate Cancer. *Prostate,* 76(9), pp. 810-822.

Rubin, M.A., Dunn, R., Strawderman, M. & Pienta, K.J., 2002. Tissue Microarray sampling strategy for prostate cancer biomarker analysis. *American Journal of Surgical Pathology,* 26(3), pp. 312-319.

Rubin, M.A., Mucci, N.R., Figurski, J., Fecko, A., Pienta, K.J. & Day, M.L., 2001. E-cadherin expression in prostate cancer: A broad survey using high-density tissue microarray technology. *Human Pathology,* 32(7), pp. 690-697.

Ruiz, C., Oeggerli, M., Germann, M., Gluderer, S., Stocker, H., Andreozzi, M., Thalmann, G.N., Cecchini, M.G., Zellweger, T., Sturm, S., Koivisto, P.A., Helin, H.J., Gelmann, E.P., Glass, A.G., Gasser, T.C., Terracciano, L.M., Bachmann, A., Wyler, S., Bubendorf, L. & Rentsch, C.A., 2012. High NRBP1 expression in prostate cancer is linked with poor clinical outcomes and increased cancer cell growth. *The Prostate,* 72(15).

Rybak, A.P., He, L.Z., Kapoor, A., Cutz, J.C. & Tang, D., 2011. Characterization of sphere-propagating cells with stem-like properties from DU145 prostate cancer cells. *Biochimica Et Biophysica Acta-Molecular Cell Research,* 1813(5), pp. 683-694.

Rye, M.B., Bertilsson, H., Drablos, F., Angelsen, A., Bathen, T.F. & Tessem, M.B., 2014. Gene signatures ESC, MYC and ERG-fusion are early markers of a potentially dangerous subtype of prostate cancer. *Bmc Medical Genomics,* 7.

Safi, R., Nelson, E.R., Chitneni, S.K., Franz, K.J., George, D.J., Zalutsky, M.R. & McDonnell, D.P., 2014. Copper Signaling Axis as a Target for Prostate Cancer Therapeutics. *Cancer Research,* 74(20), pp. 5819-5831.

Salvatori, L., Caporuscio, F., Verdina, A., Starace, G., Crispi, S., Nicotra, M.R., Russo, A., Calogero, R.A., Morgante, E., Natali, P.G., Russo, M.A. & Petrangeli, E., 2012. Cell-to-Cell Signaling Influences the Fate of Prostate Cancer Stem Cells and Their Potential to Generate More Aggressive Tumors. *Plos One,* 7(2), p. 14.

Sanchez-Beato, M., Sanchez, E., Gonzalez-Carrero, J., Morente, M., Diez, A., Sanchez-Verde, L., Martin, M.C., Cigudosa, J.C., Vidal, M. & Piris, M.A., 2006. Variability in the expression of polycomb proteins in different normal and tumoral tissues. A pilot study using tissue microarrays. *Modern Pathology,* 19(5).

Sangiorgi, E. & Capecchi, M.R., 2008. Bmi1 is expressed in vivo in intestinal stem cells. *Nature Genetics,* 40(7), pp. 915-920.

Sato, H., Minei, S., Hachiya, T., Yoshida, T. & Takimoto, Y., 2006. Fluorescence in situ hybridization analysis of c-myc amplification in stage T3N0M0 prostate cancer in Japanese patients. *International Journal of Urology,* 13(6), pp. 761-766.

Sboner, A., Demichelis, F., Calza, S., Pawitan, Y., Setlur, S.R., Hoshida, Y., Perner, S., Adami, H.O., Fall, K., Mucci, L.A., Kantoff, P.W., Stampfer, M., Andersson, S.O., Varenhorst, E., Johansson, J.E., Gerstein, M.B., Golub, T.R., Rubin, M.A. & Andren, O., 2010. Molecular sampling of prostate cancer: a dilemma for predicting disease progression. *Bmc Medical Genomics,* 3, p. 12.

Schmidt, A., Echtermeyer, F., Alozie, A., Brands, K. & Buddecke, E., 2005. Plasmin- and thrombin-accelerated shedding of syndecan-4 ectodomain generates cleavage sites at Lys(114)-Arg(115) and Lys(129)-Val(130) bonds. *Journal of Biological Chemistry,* 280(41), pp. 34441-34446.

Schroeder, A., Herrmann, A., Cherryholmes, G., Kowolik, C., Buettner, R., Pal, S., Yu, H., Mueller-Newen, G. & Jove, R., 2014. Loss of Androgen Receptor Expression Promotes a Stem-like Cell Phenotype in Prostate Cancer through STAT3 Signaling. *Cancer Research,* 74(4), pp. 1227-1237.

Shariat, S.F., Svatek, R.S., Kabbani, W., Walz, J., Lotan, Y., Karakiewicz, P.I. & Roehrborn, C.G., 2008. Prognostic value of syndecan-1 expression in patients treated with radical prostatectomy. *Bju International,* 101(2), pp. 232-237.

Sharova, L.V., Sharov, A.A., Piao, Y.L., Stagg, C.A., Amano, T., Qian, Y., Dudekula, D., Schlessinger, D. & Ko, M.S.H., 2016. Emergence of undifferentiated colonies from mouse embryonic stem cells undergoing differentiation by retinoic acid treatment. *In Vitro Cellular & Developmental Biology-Animal,* 52(5), pp. 616-624.

Sharpe, B., Beresford, M., Bowen, R., Mitchard, J. & Chalmers, A.D., 2013. Searching for Prostate Cancer Stem Cells: Markers and Methods. *Stem Cell Reviews and Reports,* 9(5), pp. 721-730.

Shimada, K., Anai, S., Fujii, T., Tanaka, N., Fujimoto, K. & Konishi, N., 2013. Syndecan-1 (CD138) contributes to prostate cancer progression by stabilizing tumour-initiating cells. *Journal of Pathology,* 231(4), pp. 495-504.

Shimada, K., Nakamura, M., De Velasco, M.A., Tanaka, M., Ouji, Y. & Konishi, N., 2009. Syndecan-1, a new target molecule involved in progression of androgen-independent prostate cancer. *Cancer Science,* 100(7), pp. 1248-1254.

Siddique, H.R., Parray, A., Tarapore, R.S., Wang, L., Mukhtar, H., Karnes, R.J., Deng, Y., Konety, B.R. & Saleem, M., 2013a. BMI1 Polycomb Group Protein Acts as a Master Switch for Growth and Death of Tumor Cells: Regulates TCF4-Transcriptional Factor-Induced BCL2 Signaling. *PloS one,* 8(5), pp. e60664-e60664.

Siddique, H.R., Parray, A., Tarapore, R.S., Wang, L., Mukhtar, H., Karnes, R.J., Deng, Y.B., Konety, B.R. & Saleem, M., 2013b. BMI1 Polycomb Group Protein Acts as a Master Switch for Growth and Death of Tumor Cells: Regulates TCF4-Transcriptional Factor-Induced BCL2 Signaling. *Plos One,* 8(5).

Siddique, H.R., Parray, A., Zhong, W., Karnes, R.J., Bergstralh, E.J., Koochekpour, S., Rhim, J.S., Konety, B.R. & Saleem, M., 2013c. BMI1, stem cell factor acting as novel serum-biomarker for Caucasian and African-American prostate cancer. *PloS one,* 8(1), p. e52993.

Siegel, R.L., Miller, K.D. & Jemal, A., 2016. Cancer Statistics, 2016. *Ca-a Cancer Journal for Clinicians,* 66(1), pp. 7-30.

Signoretti, S., Waltregny, D., Dilks, J., Isaac, B., Lin, D., Garraway, L., Yang, A., Montironi, R., McKeon, F. & Loda, M., 2000. p63 is a prostate basal cell marker and is required for prostate development. *American Journal of Pathology,* 157(6), pp. 1769-1775.

Singh, S.S., Mehedint, D.C., Ford, O.H., III, Maygarden, S.J., Ruiz, B. & Mohler, J.L., 2007. Feasibility of constructing tissue microarrays from diagnostic prostate biopsies. *Prostate,* 67(10), pp. 1011-1018.

Slovin, S.F., Wilton, A.S., Heller, G. & Scher, H.I., 2005. Time to detectable metastatic disease in patients with rising prostate-specific antigen values following surgery or radiation therapy. *Clinical Cancer Research,* 11(24), pp. 8669-8673.

Smith, M.R., Kabbinavar, F., Saad, F., Hussain, A., Gittelman, M.C., Bilhartz, D.L., Wynne, C., Murray, R., Zinner, N.R., Schulman, C., Linnartz, R., Zheng, M., Goessl, C., Hei, Y.L., Small, E.J., Cook, R. & Higano, C.S., 2005. Natural history of rising serum prostate-specific antigen in men with castrate nonmetastatic prostate cancer. *Journal of Clinical Oncology,* 23(13), pp. 2918-2925.

Sokoll, L.J., Ellis, W., Lange, P., Noteboom, J., Elliott, D.J., Deras, I.L., Blase, A., Koo, S., Sarno, M., Rittenhouse, H., Groskopf, J. & Vessella, R.L., 2008. A multicenter evaluation of the PCA3 molecular urine test: Pre-analytical effects, analytical performance, and diagnostic accuracy. *Clinica Chimica Acta,* 389(1-2), pp. 1-6.

Sooriakumaran, P., Nyberg, T., Akre, O., Haendler, L., Heus, I., Olsson, M., Carlsson, S., Roobol, M.J., Steineck, G. & Wiklund, P., 2014. Comparative effectiveness of radical prostatectomy and radiotherapy in prostate cancer: observational study of mortality outcomes. *Bmj-British Medical Journal,* 348.

Soos, G., Tsakiris, I., Szanto, J., Turzo, C., Haas, P.G. & Dezso, B., 2005. The prevalence of prostate carcinoma and its precursor in Hungary: An autopsy study. *European Urology,* 48(5), pp. 739-744.

Sorrentino, C., Musiani, P., Pompa, P., Cipollone, G. & Di Carlo, E., 2011. Androgen Deprivation Boosts Prostatic Infiltration of Cytotoxic and Regulatory T Lymphocytes and Has No Effect on Disease-Free Survival in Prostate Cancer Patients. *Clinical Cancer Research,* 17(6), pp. 1571-1581.

Sotomayor, P., Godoy, A., Smith, G.J. & Huss, W.J., 2009. Oct4A is Expressed by a Subpopulation of Prostate Neuroendocrine Cells. *Prostate,* 69(4), pp. 401-410.

Stambolic, V., Suzuki, A., de la Pompa, J.L., Brothers, G.M., Mirtsos, C., Sasaki, T., Ruland, J., Penninger, J.M., Siderovski, D.P. & Mak, T.W., 1998. Negative regulation of PKB/Akt-dependent cell survival by the tumor suppressor PTEN. *Cell,* 95(1), pp. 29-39.

Steinkamp, M.P., O'Mahony, O.A., Brogley, M., Rehman, H., LaPensee, E.W., Dhanasekaran, S., Hofer, M.D., Kuefer, R., Chinnaiyan, A., Rubin, M.A., Pienta, K.J. & Robins, D.M., 2009. Treatment-Dependent Androgen Receptor Mutations in Prostate Cancer Exploit Multiple Mechanisms to Evade Therapy. *Cancer Research,* 69(10), pp. 4434-4442.

Stephenson, A.J., Kattan, M.W., Eastham, J.A., Dotan, Z.A., Bianco, F.J., Jr., Lilja, H. & Scardino, P.T., 2006. Defining biochemical recurrence of prostate cancer after radical prostatectomy: A proposal for standardized definition. *Journal of Clinical Oncology,* 24(24), pp. 3973-3978.

Sternberg, C.N., Baskin-Bey, E.S., Watson, M., Worsfold, A., Rider, A. & Tombal, B., 2013. Treatment patterns and characteristics of European patients with castration-resistant prostate cancer. *Bmc Urology,* 13.

Stewart, M.D., Ramani, V.C. & Sanderson, R.D., 2015. Shed Syndecan-1 Translocates to the Nucleus of Cells Delivering Growth Factors and Inhibiting Histone Acetylation A NOVEL MECHANISM OF TUMOR-HOST CROSS-TALK. *Journal of Biological Chemistry,* 290(2), pp. 941-949.

Stockley, J., Villasevil, M.E.M., Nixon, C., Ahmad, I., Leung, H.Y. & Rajan, P., 2014. The RNA-binding protein hnRNPA2 regulates beta-catenin protein expression and is overexpressed in prostate cancer. *Rna Biology,* 11(6), pp. 755-765.

Storm, M.P., Kumpfmueller, B., Bone, H.K., Buchholz, M., Ripoll, Y.S., Chaudhuri, J.B., Niwa, H., Tosh, D. & Welham, M.J., 2014. Zscan4 Is Regulated by PI3-Kinase and DNA-Damaging Agents and Directly Interacts with the Transcriptional Repressors LSD1 and CtBP2 in Mouse Embryonic Stem Cells. *Plos One,* 9(3).

Su, G., Blaine, S.A., Qiao, D.H. & Friedl, A., 2007. Shedding of syndecan-1 by stromal fibroblasts stimulates human breast cancer cell proliferation via FGF2 activation. *Journal of Biological Chemistry,* 282(20), pp. 14906-14915.

Sun, Y. & Goodison, S., 2009. Optimizing Molecular Signatures for Predicting Prostate Cancer Recurrence. *Prostate,* 69(10), pp. 1119-1127.

Sung, Y.Y. & Cheung, E., 2014. Androgen receptor co-regulatory networks in castration-resistant prostate cancer. *Endocrine-Related Cancer,* 21(1), pp. R1-R11.

Supek, F., Bosnjak, M., Skunca, N. & Smuc, T., 2011. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *Plos One,* 6(7).

Szarvas, T., Reis, H., vom Dorp, F., Tschirdewahn, S., Niedworok, C., Nyirady, P., Schmid, K.W., Rubben, H. & Kovalszky, I., 2016a. Soluble syndecan-1 (SDC1) serum level as an independent pre-operative predictor of cancer-specific survival in prostate cancer. *Prostate,* 76(11), pp. 977-985.

Szarvas, T., Reis, H., vom Dorp, F., Tschirdewahn, S., Niedworok, C., Nyirady, P., Schmid, K.W., Rübben, H. & Kovalszky, I., 2016b. Soluble syndecan-1 (SDC1) serum level as an independent pre-operative predictor of cancer-specific survival in prostate cancer. *The Prostate,* [Epub ahead of Print].

Szatmari, T., Mundt, F., Heidari-Hamedani, G., Zong, F., Ferolla, E., Alexeyenko, A., Hjerpe, A. & Dobra, K., 2012. Novel Genes and Pathways Modulated by Syndecan-1: Implications for the Proliferation and Cell-Cycle Regulation of Malignant Mesothelioma Cells. *Plos One,* 7(10).

Tagliaferri, D., De Angelis, M.T., Russo, N.A., Marotta, M., Ceccarelli, M., Del Vecchio, L., De Felice, M. & Falco, G., 2016. Retinoic Acid Specifically Enhances Embryonic Stem Cell Metastate Marked by Zscan4. *Plos One,* 11(2).

Tarasov, V., Jung, P., Verdoodt, B., Lodygin, D., Epanchintsev, A., Menssen, A., Meister, G. & Hermeking, H., 2007. Differential regulation of microRNAs by p53 revealed by massively parallel Sequencing - miR-34a is a p53 target that induces apoptosis and G(1)-arrest. *Cell Cycle,* 6(13), pp. 1586-1593.

Taylor, B.S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y.H., Carver, B.S., Arora, V.K., Kaushik, P., Cerami, E., Reva, B., Antipin, Y., Mitsiades, N., Landers, T., Dolgalev, I., Major, J.E., Wilson, M., Socci, N.D., Lash, A.E., Heguy, A., Eastham, J.A., Scher, H.I., Reuter, V.E., Scardino, P.T., Sander, C., Sawyers, C.L. & Gerald, W.L., 2010. Integrative Genomic Profiling of Human Prostate Cancer. *Cancer Cell,* 18(1), pp. 11-22.

Taylor, R.A. & Risbridger, G.P., 2008. The path toward identifying prostatic stem cells. *Differentiation,* 76(6), pp. 671-681.

Taylor, R.A., Toivanen, R., Frydenberg, M., Pedersen, J., Harewood, L., Collins, A.T., Maitland, N.J., Risbridger, G.P. & Australian Prostate, C., 2012. Human Epithelial Basal Cells Are Cells of Origin of Prostate Cancer, Independent of CD133 Status. *Stem Cells,* 30(6), pp. 1087-1096.

Tennstedt, P., Koester, P., Bruechmann, A., Mirlacher, M., Haese, A., Steuber, T., Sauter, G., Huland, H., Graefen, M., Schlomm, T., Minner, S. & Simon, R., 2012. The impact of the number of cores on tissue microarray studies investigating prostate cancer biomarkers. *International Journal of Oncology,* 40(1), pp. 261-268.

Thompson, I.M., Pauler, D.K., Goodman, P.J., Tangen, C.M., Lucia, M.S., Parnes, H.L., Minasian, L.M., Ford, L.G., Lippman, S.M., Crawford, E.D., Crowley, J.J. & Coltman, C.A., 2004. Prevalence of prostate cancer among men with a prostate-specific antigen level <= 4.0 ng per milliliter. *New England Journal of Medicine,* 350(22), pp. 2239-2246.

Thorvaldsdottir, H., Robinson, J.T. & Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics,* 14(2), pp. 178-192.

Tirino, V., Desiderio, V., Paino, F., De Rosa, A., Papaccio, F., La Noce, M., Laino, L., De Francesco, F. & Papaccio, G., 2013. Cancer stem cells in solid tumors: an overview and new approaches for their isolation and characterization. *Faseb Journal,* 27(1), pp. 13-24.

Tolonen, T.T., Kujala, P.M., Laurila, M., Tirkkonen, M., Ilvesaro, J., Tuominen, V.J., Tammela, T.L.J. & Isola, J., 2011. Routine dual-color immunostaining with a 3-antibody cocktail improves the detection of small cancers in prostate needle biopsies. *Human Pathology,* 42(11), pp. 1635-1642.

Tomlins, S.A., Alshalalfa, M., Davicioni, E., Erho, N., Yousefi, K., Zhao, S., Haddad, Z., Den, R.B., Dicker, A.P., Trock, B.J., DeMarzo, A.M., Ross, A.E., Schaeffer, E.M., Klein, E.A., Magi-Galluzzi, C., Karnes, R.J., Jenkins, R.B. & Feng, F.Y., 2015. Characterization of 1577 Primary Prostate Cancers Reveals Novel Biological and Clinicopathologic Insights into Molecular Subtypes. *European Urology,* 68(4), pp. 555-567.

Tomlins, S.A., Day, J.R., Lonigro, R.J., Hovelson, D.H., Siddiqui, J., Kunju, L.P., Dunn, R.L., Meyer, S., Hodge, P., Groskopf, J., Wei, J.T. & Chinnaiyan, A.M., 2016. Urine TMPRSS2:ERG Plus PCA3 for Individualized Prostate Cancer Risk Assessment. *European Urology,* 70(1), pp. 45-53.

Tomlins, S.A., Laxman, B., Varambally, S., Cao, X., Yu, J., Helgeson, B.E., Cao, Q., Prensner, J.R., Rubin, M.A., Shah, R.B., Mehra, R. & Chinnaiyan, A.M., 2008. Role of the TMPRSS2-ERG Gene Fusion in Prostate Cancer. *Neoplasia.* pp. 177-88.

Tomlins, S.A., Mehra, R., Rhodes, D.R., Cao, X.H., Wang, L., Dhanasekaran, S.M., Kalyana-Sundaram, S., Wei, J.T., Rubin, M.A., Pienta, K.J., Shah, R.B. & Chinnaiyan, A.M., 2007. Integrative molecular concept modeling of prostate cancer progression. *Nature Genetics,* 39(1), pp. 41-51.

Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X.H., Tchinda, J., Kuefer, R., Lee, C., Montie, J.E., Shah, R.B., Pienta, K.J., Rubin, M.A. & Chinnaiyan, A.M., 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science,* 310(5748), pp. 644-648.

Torabi-Nezhad, S., Malekmakan, L., Mashayekhi, M. & Daneshian, A., 2016. Histopathological features of intra-ductal carcinoma of prostatic and high grade prostatic intraepithelialneoplasia and correlation with PTEN and P63. *Prostate,* 76(4), pp. 394-401.

Totten, R.S., Heinemann, M.W., Hudson, P.B., Sproul, E.E. & Stout, A.P., 1953. MICROSCOPIC DIFFERENTIAL DIAGNOSIS OF LATENT CARCINOMA OF PROSTATE. *Ama Archives of Pathology,* 55(2), pp. 131-141.

Tsai, C.Y., Finley, J.C., Ali, S.S., Patel, H.H. & Howell, S.B., 2012. Copper influx transporter 1 is required for FGF, PDGF and EGF-induced MAPK signaling. *Biochemical Pharmacology,* 84(8), pp. 1007-1013.

Tsihlias, J., Kapusta, L.R., DeBoer, G., Morava-Protzner, I., Zbieranowski, I., Bhattacharya, N., Catzavelos, G.C., Klotz, L.H. & Slingerland, J.M., 1998. Loss of cyclin-dependent kinase inhibitor

p27(Kip1) is a novel prognostic factor in localized human prostate adenocarcinoma. *Cancer Research,* 58(3), pp. 542-548.

Tsujimura, A., Koikawa, Y., Salm, S., Takao, T., Coetzee, S., Moscatelli, D., Shapiro, E., Lepor, H., Sun, T.T. & Wilson, E.L., 2002. Proximal location of mouse prostate epithelial stem cells: a model of prostatic homeostasis. *Journal of Cell Biology,* 157(7), pp. 1257-1265.

Turay, D., Khan, S., Osterman, C.J.D., Curtis, M.P., Khaira, B., Neidigh, J.W., Mirshahidi, S., Casiano, C.A. & Wall, N.R., 2016. Proteomic Profiling of Serum-Derived Exosomes from Ethnically Diverse Prostate Cancer Patients. *Cancer Investigation,* 34(1), pp. 1-11.

Udager, A.M., DeMarzo, A.M., Shi, Y., Hicks, J.L., Cao, X.H., Siddiqui, J., Jiang, H., Chinnaiyan, A.M. & Mehra, R., 2016. Concurrent Nuclear ERG and MYC Protein Overexpression Defines a Subset of Locally Advanced Prostate Cancer: Potential Opportunities for Synergistic Targeted Therapeutics. *Prostate,* 76(9), pp. 845-853.

Ugolkov, A.V., Eisengart, L.J., Luan, C.Y. & Yang, X.M.J., 2011. Expression Analysis of Putative Stem Cell Markers in Human Benign and Malignant Prostate. *Prostate,* 71(1), pp. 18-25.

Ulz, P., Belic, J., Graf, R., Auer, M., Lafer, I., Fischereder, K., Webersinke, G., Pummer, K., Augustin, H., Pichler, M., Hoefler, G., Bauernhofer, T., Geigl, J.B., Heitzer, E. & Speicher, M.R., 2016. Whole-genome plasma sequencing reveals focal amplifications as a driving force in metastatic prostate cancer. *Nature Communications,* 7.

van den Hoogen, C., van der Horst, G., Cheung, H., Buijs, J.T., Lippitt, J.M., Guzman-Ramirez, N., Hamdy, F.C., Eaton, C.L., Thalmann, G.N., Cecchini, M.G., Pelger, R.C.M. & van der Pluijm, G., 2010. High Aldehyde Dehydrogenase Activity Identifies Tumor-Initiating and Metastasis-Initiating Cells in Human Prostate Cancer. *Cancer Research,* 70(12), pp. 5163-5173.

van den Hoogen, C., van der Horst, G., Cheung, H., Buijs, J.T., Pelger, R.C.M. & van der Pluijm, G., 2011. The aldehyde dehydrogenase enzyme 7A1 is functionally involved in prostate cancer bone metastasis. *Clinical & Experimental Metastasis,* 28(7).

van der Horst, G., van den Hoogen, C., Buijs, J.T., Cheung, H., Bloys, H., Pelger, R.C.M., Lorenzon, G., Heckmann, B., Feyen, J., Pujuguet, P., Blanque, R., Clement-Lacroix, P. & van der Pluijm, G., 2011. Targeting of alpha(v)-Integrins in Stem/Progenitor Cells and Supportive Microenvironment Impairs Bone Metastasis in Human Prostate Cancer. *Neoplasia,* 13(6).

van Leenders, G., Dijkman, H., Hulsbergen-van de Kaa, C., Ruiter, D. & Schalken, J., 2000. Demonstration of intermediate cells during human prostate epithelial differentiation in situ and in vitro using triple-staining confocal scanning microscopy. *Laboratory Investigation,* 80(8), pp. 1251-1258.

van Leenders, G.J.L.H., Dukers, D., Hessels, D., van den Kieboom, S.W.M., Hulsbergen, C.A., Witjes, J.A., Otte, A.P., Meijer, C.J. & Raaphorst, F.M., 2007. Polycomb-group oncogenes EZH2, BMI1, and RING1 are overexpressed in prostate cancer with adverse pathologic and clinical features. *European Urology,* 52(2).

Varambally, S., Dhanasekaran, S.M., Zhou, M., Barrette, T.R., Kumar-Sinha, C., Sanda, M.G., Ghosh, D., Pienta, K.J., Sewalt, R., Otte, A.P., Rubin, M.A. & Chinnaiyan, A.M., 2002. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature,* 419(6907), pp. 624-629.

Varma, M. & Jasani, B., 2005. Diagnostic utility of immunohistochemistry in morphologically difficult prostate cancer: review of current literature. *Histopathology,* 47(1), pp. 1-16.

Varma, M., Lee, M.W., Tamboli, P., Zarbo, R.J., Jimenez, R.E., Salles, P.G.O. & Amin, M.B., 2002. Morphologic criteria for the diagnosis of prostatic adenocarcinoma in needle biopsy specimens - A study of 250 consecutive cases in a routine surgical pathology practice. *Archives of Pathology & Laboratory Medicine,* 126(5), pp. 554-561.

Vesuna, F., Lisok, A., Kimble, B. & Raman, V., 2009. Twist Modulates Breast Cancer Stem Cells by Transcriptional Regulation of CD24 Expression. *Neoplasia,* 11(12).

Viegas, M.S., Martins, T.C., Seco, F. & do Carmo, A., 2007. An improved and cost-effective methodology for the reduction of autofluorescence in direct immunofluorescence studies on formalin-fixed paraffin-embedded tissues. *European Journal of Histochemistry,* 51(1), pp. 59-66.

Visakorpi, T., Hyytinen, E., Koivisto, P., Tanner, M., Keinanen, R., Palmberg, C., Palotie, A., Tammela, T., Isola, J. & Kallioniemi, O.P., 1995. IN-VIVO AMPLIFICATION OF THE ANDROGEN RECEPTOR GENE AND PROGRESSION OF HUMAN PROSTATE-CANCER. *Nature Genetics,* 9(4), pp. 401-406.

Visvader, J.E. & Lindeman, G.J., 2008. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature Reviews Cancer,* 8(10), pp. 755-768.

Voyvodic, P.L., Min, D., Liu, R., Williams, E., Chitalia, V., Dunn, A.K. & Baker, A.B., 2014. Loss of Syndecan-1 Induces a Pro-inflammatory Phenotype in Endothelial Cells with a Dysregulated Response to Atheroprotective Flow*. *Journal of Biological Chemistry,* 289(14), pp. 9547-9559.

Vuoriluoto, K., Hognas, G., Meller, P., Lehti, K. & Ivaska, J., 2011. Syndecan-1 and-4 differentially regulate oncogenic K-ras dependent cell invasion into collagen through alpha 2 beta 1 integrin and MT1-MMP. *Matrix Biology,* 30(3), pp. 207-217.

Vuoriluoto, K., Jokinen, J., Kallio, K., Salmivirta, M., Heino, J. & Ivaska, J., 2008. Syndecan-1 supports integrin alpha 2 beta 1-mediated adhesion to collagen. *Experimental Cell Research,* 314(18), pp. 3369-3381.

Wang, H., Huang, M., Zhang, D.Y. & Zhang, F., 2011. Global Profiling of Signaling Networks: Study of Breast Cancer Stem Cells and Potential Regulation. *Oncologist,* 16(7).

Wang, H.Y., Xu, Y.Y., Fang, Z., Chen, S., Balk, S.P. & Yuan, X., 2012a. Doxycycline Regulated Induction of AKT in Murine Prostate Drives Proliferation Independently of p27 Cyclin Dependent Kinase Inhibitor Downregulation. *Plos One,* 7(7).

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W. & Orkin, S.H., 2006. A protein interaction network for pluripotency of embryonic stem cells. *Nature,* 444(7117).

Wang, Q.J., Yang, L.D., Alexander, C. & Temple, S., 2012b. The Niche Factor Syndecan-1 Regulates the Maintenance and Proliferation of Neural Progenitor Cells during Mammalian Cortical Development. *Plos One,* 7(8).

Wang, X., Kruithof-de Julio, M., Economides, K.D., Walker, D., Yu, H.L., Halili, M.V., Hu, Y.P., Price, S.M., Abate-Shen, C. & Shen, M.M., 2009. A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature,* 461(7263), pp. 495-U61.

Wang, Z.A., Mitrofanova, A., Bergren, S.K., Abate-Shen, C., Cardiff, R.D., Califano, A. & Shen, M.M., 2013. Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. *Nature cell biology,* 15(3), pp. 274-83.

Wei, W., Leng, J.Y., Shao, H.X. & Wang, W.D., 2015. High PCA3 scores in urine correlate with poor-prognosis factors in prostate cancer patients. *International Journal of Clinical and Experimental Medicine,* 8(9), pp. 16606-16612.

Wen, S.M., Tian, J., Niu, Y.J., Li, L., Yeh, S.Y. & Chang, C.S., 2016. ASC-J9 (R), and not Casodex or Enzalutamide, suppresses prostate cancer stem/progenitor cell invasion via altering the EZH2-STAT3 signals. *Cancer Letters,* 376(2), pp. 377-386.

Whiting, P.F., Moore, T.H.M., Jameson, C.M., Davies, P., Rowlands, M.A., Burke, M., Beynon, R., Savovic, J. & Donovan, J.L., 2016. Symptomatic and quality-of-life outcomes after treatment for clinically localised prostate cancer: a systematic review. *Bju International,* 118(2), pp. 193-204.

Wu, W.W.H., Wong, J.P., Kast, J. & Molday, R.S., 2005. RS1, a discoidin domain-containing retinal cell adhesion protein associated with X-linked retinoschisis, exists as a novel disulfide-linked octamer. *Journal of Biological Chemistry,* 280(11), pp. 10721-10730.

Wuidart, A., Ousset, M., Rulands, S., Simons, B.D., Van Keymeulen, A. & Blanpain, C., 2016. Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes & Development,* 30(11), pp. 1261-1277.

Wyatt, A.W., Mo, F., Wang, K., McConeghy, B., Brahmbhatt, S., Jong, L.N., Mitchell, D.M., Johnston, R.L., Haegert, A., Li, E., Liew, J., Yeung, J., Shrestha, R., Lapuk, A.V., McPherson, A., Shukin, R., Bell, R.H., Anderson, S., Bishop, J., Hurtado-Coll, A., Xiao, H., Chinnaiyan, A.M., Mehra, R., Lin, D., Wang, Y.Z., Fazli, L., Gleave, M.E., Volik, S.V. & Collins, C.C., 2014. Heterogeneity in the inter-tumor transcriptome of high risk prostate cancer. *Genome Biology,* 15(8).

Xiao, J.T., Angsana, J., Wen, J., Smith, S.V., Park, P.W., Ford, M.L., Haller, C.A. & Chaikof, E.L., 2012. Syndecan-1 Displays a Protective Role in Aortic Aneurysm Formation by Modulating T Cell-Mediated Responses. *Arteriosclerosis Thrombosis and Vascular Biology,* 32(2), pp. 386-U519.

Xie, Y., Xu, K., Linn, D.E., Yang, X., Guo, Z., Shimelis, H., Nakanishi, T., Ross, D.D., Chen, H., Fazli, L., Gleave, M.E. & Qiu, Y., 2008. The 44-kDa pim-1 kinase phosphorylates BCRP/ABCG2 and thereby promotes its multimerization and drug-resistant activity in human prostate cancer cells. *Journal of Biological Chemistry,* 283(6), pp. 3349-3356.

Xu, F., Xiang, H., Jiang, R., Dong, F. & Sui, R., 2011. Phenotypic expression of X-linked retinoschisis in Chinese families with mutations in the RS1 gene. *Documenta Ophthalmologica,* 123(1), pp. 21-27.

Yang, D.-D., Cui, B.-B., Sun, L.-y., Zheng, H.-q., Huang, Q., Tong, J.-X. & Zhang, Q.-F., 2011. The Co-expression of USP22 and BMI-1 May Promote Cancer Progression and Predict Therapy Failure in Gastric Carcinoma. *Cell Biochemistry and Biophysics,* 61(3).

Yang, X.J., Lin, X.J., Zhong, X.M., Kaur, S., Li, N., Liang, S., Lassus, H., Wang, L.P., Katsaros, D., Montone, K., Zhao, X., Zhang, Y.C., Butzow, R., Coukos, G. & Zhang, L., 2010. Double-Negative Feedback Loop between Reprogramming Factor LIN28 and microRNA let-7 Regulates Aldehyde Dehydrogenase 1-Positive Cancer Stem Cells. *Cancer Research,* 70(22), pp. 9463-9472.

Yildiz-Sezer, S., Verdorfer, I., Schafer, G., Rogatsch, H., Bartsch, G. & Mikuz, G., 2006. Assessment of aberrations on chromosome 8 in prostatic atrophy. *Bju International,* 98(1), pp. 184-188.

Yin, A.H., Miraglia, S., Zanjani, E.D., AlmeidaPorada, G., Ogawa, M., Leary, A.G., Olweus, J., Kearney, J. & Buck, D.W., 1997. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood,* 90(12), pp. 5002-5012.

Yin, H. & Glass, J., 2011. The Phenotypic Radiation Resistance of CD44(+)/CD24(-or) (low) Breast Cancer Cells Is Mediated through the Enhanced Activation of ATM Signaling. *Plos One,* 6(9), p. 11.

Younesi, E., Toldo, L., Muller, B., Friedrich, C.M., Novac, N., Scheer, A., Hofmann-Apitius, M. & Fluck, J., 2012. Mining biomarker information in biomedical literature. *Bmc Medical Informatics and Decision Making,* 12.

Yousef, G.M. & Diamandis, E.P., 2001. The new human tissue kallikrein gene family: Structure, function, and association to disease. *Endocrine Reviews,* 22(2), pp. 184-204.

Yu, L., Toriseva, M., Tuomala, M., Seikkula, H., Elo, T., Tuomela, J., Kallajoki, M., Mirtti, T., Taimen, P., Bostrom, P.J., Alanen, K., Nurmi, M., Nees, M. & Harkonen, P., 2016. Increased expression of fibroblast growth factor 13 in prostate cancer is associated with shortened time to biochemical recurrence after radical prostatectomy. *International Journal of Cancer,* 139(1), pp. 140-152.

Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S.-L., Stagg, C.A., Hoang, H.G., Yang, H.-T., Indig, F.E., Wersto, R.P. & Ko, M.S.H., 2010. Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature,* 464(7290).

Zellweger, T., Ninck, C.H., Mirlacher, M., Annefeld, M., Glass, A.G., Gasser, T.C., Mihatsch, M.J., Gelmann, E.P. & Bubendorf, L., 2003. Tissue microarray analysis reveals prognostic significance of syndecan-1 expression in pro-state cancer. *Prostate,* 55(1), pp. 20-29.

Zhang, K., Fowler, M., Glass, J. & Yin, H., 2014. Activated 5 ' Flanking Region ofNANOGP8 in a Self-Renewal Environment isAssociated with Increased Sphere Formation and Tumor Growth of Prostate Cancer Cells. *Prostate,* 74(4), pp. 381-394.

Zhang, L.L., Jiao, M., Li, L., Wu, D.P., Wu, K.J., Li, X., Zhu, G.D., Dang, Q., Wang, X.Y., Hsieh, J.T. & He, D.L., 2012a. Tumorspheres derived from prostate cancer cells possess chemoresistant and cancer stem cell properties. *Journal of Cancer Research and Clinical Oncology,* 138(4), pp. 675-686.

Zhang, L.L., Li, L., Jiao, M., Wu, D.P., Wu, K.J., Li, X., Zhu, G.D., Yang, L., Wang, X.Y., Hsieh, J.T. & He, D.L., 2012b. Genistein inhibits the sternness properties of prostate cancer cells through targeting Hedgehog-Gli1 pathway. *Cancer Letters,* 323(1), pp. 48-57.

Zhang, Q., Dan, J.M., Wang, H., Guo, R.P., Mao, J., Fu, H.F., Wei, X.W. & Liu, L., 2016. Tcstv1 and Tcstv3 elongate telomeres of mouse ES cells. *Scientific Reports,* 6.

Zhang, W.S., Zhu, J., Efferson, C.L., Ware, C., Tammam, J., Angagaw, M., Laskey, J., Bettano, K.A., Kasibhatla, S., Reilly, J.F., Sur, C. & Majumder, P.K., 2009. Inhibition of Tumor Growth Progression by Antiandrogens and mTOR Inhibitor in a Pten-Deficient Mouse Model of Prostate Cancer. *Cancer Research,* 69(18), pp. 7466-7472.

Zhang, Y., Yao, L., Zhang, X., Ji, H., Wang, L., Sun, S. & Pang, D., 2011. Elevated expression of USP22 in correlation with poor prognosis in patients with invasive breast cancer. *Journal of Cancer Research and Clinical Oncology,* 137(8).

Zhou, M., Aydin, H., Kanane, H. & Epstein, J.I., 2004. How often does alpha-methylacyl-CoA-racemase contribute to resolving an atypical diagnosis on prostate needle biopsy beyond that provided by basal cell markers? *American Journal of Surgical Pathology,* 28(2), pp. 239-243.

Zhou, S., Schuetz, J.D., Bunting, K.D., Colapietro, A.M., Sampath, J., Morris, J.J., Lagutina, I., Grosveld, G.C., Osawa, M., Nakauchi, H. & Sorrentino, B.P., 2001. The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nature Medicine,* 7(9), pp. 1028-1034.

Zong, F., Fthenou, E., Mundt, F., Szatmari, T., Kovalszky, I., Szilak, L., Brodin, D., Tzanakakis, G., Hjerpe, A. & Dobra, K., 2011. Specific Syndecan-1 Domains Regulate Mesenchymal Tumor Cell Adhesion, Motility and Migration. *Plos One,* 6(6).

Zong, F., Fthenou, E., Wolmer, N., Hollosi, P., Kovalszky, I., Szilak, L., Mogler, C., Nilsonne, G., Tzanakakis, G. & Dobra, K., 2009. Syndecan-1 and FGF-2, but Not FGF Receptor-1, Share a Common Transport Route and Co-Localize with Heparanase in the Nuclei of Mesenchymal Tumor Cells. *Plos One,* 4(10).

Zou, J.X., Zhong, Z.Y., Shi, X.B., Tepper, C.G., White, R.W.D., Kung, H.J. & Chen, H.W., 2006. ACTR/AIB1/SRC-3 and androgen receptor control prostate cancer cell proliferation and tumor growth through direct control of cell cycle genes. *Prostate,* 66(14), pp. 1474-1486.

8 APPENDICES

APPENDIX 1: CODE FOR EXTREMES OF EXPRESSION ANALYSIS AND SUBSEQUENT GO ENRICHMENT ANALYSIS

```
# Extremes of Expression Analysis R Code Functions
# Segregates data on 10th/90th percentiles of expression
# Conducts a differential expression analysis on these groups
# Attempts to determine functions co-regulated with genes
# of interest in microarray expression data.
# Author: Benjamin Peter Sharpe
#####################################################################
#Internal functions for dichotomising the data on 10th/90th percentiles
#IDs for a gene must be looked up manually using the platform information or by calling fData() on the ExpressionSet to extract feature info.
#Eset is expression data that can be generated by calling exprs() on an object of class ExpressionSet, producing a 2d array of intensities labelled by sample ID and probe ID.

Extreme_lower<-function(Eset, Probe){lowertruncated<-Eset[,Eset[Probe,]<quantile(Eset[Probe,], 1/10)] #Identifies samples in the lower 10th percentile
        return(colnames(lowertruncated))}
Extreme_upper<-function(Eset, Probe){uppertruncated<-Eset[,Eset[Probe,]>quantile(Eset[Probe,], 9/10)] #Identifies samples in the upper 90th percentile
        return(colnames(uppertruncated))}
Function.LowTruthsAsOne<-function(q){if(q==TRUE) 1 else 0}
Function.HighTruthsAsTwo<-function(y){if(y==TRUE) 2 else 0}
Function.RecogniseNumbersAsFactors<-function(z){if(z==0) "MID" else( #Function for recognising samples as low, medium or high categories based on above functions.
```

```
                    if(z==1) "LOW" else "HIGH")}

Factors.Extreme<-function(Eset, Probe){ #Returns a list of samples labelled with "LOW", "MID", or "HIGH" based on intensity values of probe.

        Lower.Matrix<-colnames(Eset) %in% Extreme_lower(Eset, Probe)

        Upper.Matrix<-colnames(Eset) %in% Extreme_upper(Eset, Probe)

        LowFactorComponent<-unlist(lapply(Lower.Matrix, Function.LowTruthsAsOne))

        HighFactorComponent<-unlist(lapply(Upper.Matrix, Function.HighTruthsAsTwo))

        Numeric.Matrix<-LowFactorComponent + HighFactorComponent

        Factor.List<-as.factor(unlist(lapply(Numeric.Matrix, Function.RecogniseNumbersAsFactors)))

        return(Factor.List)

}


#Internal functions for colour assignment of heatmaps in R based on factors passed from Factors.Extreme

#Assigns blue colour label to 'low' samples, yellow colour label to 'high' samples and optional grey colour to 'mid' samples (if boolUseMidSamples is set to true it will heatmap intermediate samples also)

Color.assignment<-function(x, boolUseMidSamples=c(FALSE, TRUE)){if(boolUseMidSamples==FALSE){if(x=="LOW") "BLUE" else(if(x=="HIGH") "YELLOW")} else {if(x=="LOW") "BLUE" else(if(x=="HIGH") "YELLOW" else "GRAY")}}

Color.labels<-function(Factor, boolUseMidSamples){

        unlist(lapply(Factor, Color.assignment, boolUseMidSamples=boolUseMidSamples))}


#User-run function.
```

#Eset must be extracted by exprs() on ExpressionSet, Probe manually identified as above. Intermediate samples (10%<x<90%) are plotted if boolPlotMidSamples is set to true, default false.

##Function plots a heatmap contrasting top and bottom 10 percentiles with one another, conducts differential gene expression analysis on these two groups and writes a csv file with the results.

#boolPlotSubset will plot only the top 50 differentially expressed genes (where they are many) - default true. Otherwise, all are plotted. ResultsFileName writes a csv file to the specified name.

#This function will work with Agilent Whole Genome Microarray platforms such as GSE35988. The structure of fData is platform-dependent and will need code adjustment for different platforms.

```
ExtremesOfExpression.Agilent<-function(ExpressionSet, Eset, Probe,
boolPlotMidSamples=c(FALSE, TRUE), boolPlotSubset=c(TRUE, FALSE),
ResultsFileName){
        require(Biobase)
        require(limma)
        require(gplots)
        AnnotationLookup<-fData(ExpressionSet)[,c("ID", "GENE_SYMBOL",
"ENSEMBL_ID")]
        AnnotationLookup<-
AnnotationLookup[!(AnnotationLookup$GENE_SYMBOL==""),] #Removing all
unannotated probes
        AnnotationLookup<-AnnotationLookup[AnnotationLookup[,"ID"] %in%
rownames(Eset),] #Ensuring a match between Annotations and probes on the Eset
        f<-Factors.Extreme(Eset, Probe)
        phenoData(ExpressionSet)$CATEGORY<-f        #Make a new phenotypic
data annotation to store Low/mid/high expression categorisation.
        Eset<-Eset[,!(ExpressionSet@phenoData$CATEGORY=="MID")]
        #Filter out samples labelled with "MID" from the DGE analysis.
```

```
f<-factor(f[f %in% c("LOW", "HIGH")])

Eset<-Eset[rownames(Eset) %in% AnnotationLookup[,"ID"],]

design<-model.matrix(~f)

fit<-eBayes(lmFit(Eset, design))      #Fit linear model to the data

TopTable<-topTable(fit, coef=2, number=1000) #Reports top 1000 ranked genes from the linear model fit

TopTable<-merge(TopTable, AnnotationLookup, by="row.names")
#Merge ranked genes with feature data including gene symbols.

TopTable<-TopTable[,c("Row.names", "logFC", "AveExpr", "t", "P.Value", "adj.P.Val", "B", "GENE_SYMBOL", "ENSEMBL_ID")]

TopTable<-TopTable[(TopTable$adj.P.Val<0.05),] #Only write results significant at the level of p<0.05, P-value adjusted for multiple testing.

if(boolPlotMidSamples==FALSE){

    Eset_Extremes<-Eset[,f %in% c("LOW", "HIGH")]}

else{Eset_Extremes<-Eset}

rownames(Eset_Extremes)<-AnnotationLookup[,"GENE_SYMBOL"]

selected<-p.adjust(fit$p.value[,2], method="holm") < 0.05#Use Holm-Bonferonni method to correct p-value for multiple comparisons testing.

Eset_Extremes_Selected<-Eset_Extremes[selected,]

if(boolPlotSubset==TRUE){

    Eset_Extremes_Selected<-Eset_Extremes_Selected[1:50,]} #Plot only top 50 ranked genes if boolPlotSubset=TRUE

heatmap.2(Eset_Extremes_Selected, col=redgreen(75), scale="row", ColSideColors = Color.labels(f, boolPlotMidSamples), key=TRUE, symkey=FALSE, density.info="none", trace="none", cexRow=0.8, cexCol= 0.8)

write.csv(TopTable, file=ResultsFileName)

}
```

```
#Identical function to ExtremesOfExpression.Agilent, but with fData structure adapted to Illumina expression beadchip platforms such as those used by GSE70768.

ExtremesOfExpression.Illumina<-function(ExpressionSet, Eset, Probe, boolPlotMidSamples=c(FALSE, TRUE), boolPlotSubset=c(TRUE, FALSE), ResultsFileName){

        require(Biobase)

        require(limma)

        require(gplots)

        AnnotationLookup<-fData(ExpressionSet)[,c("ID", "Gene symbol")]

        AnnotationLookup<-AnnotationLookup[!(AnnotationLookup$"Gene symbol"==""),] #Removing all unannotated probes

        AnnotationLookup<-AnnotationLookup[AnnotationLookup[,"ID"] %in% rownames(Eset),] #Ensuring a match between Annotations and probes on the Eset

        f<-Factors.Extreme(Eset, Probe)

        phenoData(ExpressionSet)$CATEGORY<-f

        Eset<-Eset[,!(ExpressionSet@phenoData$CATEGORY=="MID")]

        f<-factor(f[f %in% c("LOW", "HIGH")])

        Eset<-Eset[rownames(Eset) %in% AnnotationLookup[,"ID"],]

        design<-model.matrix(~f)

        fit<-eBayes(lmFit(Eset, design))

        TopTable<-topTable(fit, coef=2, number=1000)

        TopTable<-merge(TopTable, AnnotationLookup, by="row.names")

        TopTable<-TopTable[,c("Row.names", "logFC", "AveExpr", "t", "P.Value", "adj.P.Val", "B", "Gene symbol")]

        TopTable<-TopTable[(TopTable$adj.P.Val<0.05),] #Only write results significant at the level of p<0.05, P-value adjusted for multiple testing.

        if(boolPlotMidSamples==FALSE){

                Eset_Extremes<-Eset[,f %in% c("LOW", "HIGH")]}
```

```
        else{Eset_Extremes<-Eset}
        rownames(Eset_Extremes)<-AnnotationLookup[,"Gene symbol"]
        selected<-p.adjust(fit$p.value[,2], method="holm") < 0.05
        Eset_Extremes_Selected<-Eset_Extremes[selected,]
        if(boolPlotSubset==TRUE){
            Eset_Extremes_Selected<-Eset_Extremes_Selected[1:50,]}
        heatmap.2(Eset_Extremes_Selected, col=redgreen(75), scale="row",
ColSideColors = Color.labels(f, boolPlotMidSamples), key=TRUE, symkey=FALSE,
density.info="none", trace="none", cexRow=0.8, cexCol= 0.8)
        write.csv(TopTable, file=ResultsFileName)
}

#User function designed to generate list of GO annotations for each feature on the
platform used. i.e. generates "Gene Universe". Used for Agilent platforms.
makeGOAnnotations.Agilent<-function(ExpressionSet){
        require(biomaRt)
        GeneUniverse<-fData(ExpressionSet)[,c("ID", "GENE_SYMBOL")]
        GeneUniverse<-GeneUniverse[!(GeneUniverse$GENE_SYMBOL == ""),]
        GeneUniverse<-unique(GeneUniverse$GENE_SYMBOL)
        mart<-useMart(biomart="ensembl", dataset="hsapiens_gene_ensembl")
        GOAnnotations<-getBM(attributes=c("hgnc_symbol", "go_id"),
filter="hgnc_symbol", values=GeneUniverse, mart=mart) #Use BiomaRt to get a list of
gene symbols and GO IDs which correspond to the gene symbols in our gene universe.
        GOAnnotations<-aggregate(GOAnnotations$go_id,
by=list(GOAnnotations$hgnc_symbol), paste) #Process the data so that it is in the
correct format to be an annotation file for topGO.
        row.names(GOAnnotations)<-GOAnnotations[,1]
        GOAnnotations<-cbind(row.names(GOAnnotations), GOAnnotations$x)
```

```
	row.names(GOAnnotations)<-GOAnnotations[,1]
	GOAnnotations<-GOAnnotations[,2]
	return(GOAnnotations)
}

#Illumina beadchip platforms have a different fData structure. Adjusted function of makeGOAnnotations to suit.
makeGOAnnotationsIllumina<-function(ExpressionSet){
	require(biomaRt)
	GeneUniverse<-fData(ExpressionSet)[,c("ID", "Gene symbol")]
	GeneUniverse<-GeneUniverse[!(GeneUniverse$"Gene symbol" == ""),]
	GeneUniverse<-unique(GeneUniverse$"Gene symbol")
	mart<-useMart(biomart="ensembl", dataset="hsapiens_gene_ensembl")
	GOAnnotations<-getBM(attributes=c("hgnc_symbol", "go_id"), filter="hgnc_symbol", values=GeneUniverse, mart=mart) #Use BiomaRt to get a list of gene symbols and GO IDs which correspond to the gene symbols in our gene universe.
	GOAnnotations<-aggregate(GOAnnotations$go_id, by=list(GOAnnotations$hgnc_symbol), paste) #Process the data so that it is in the correct format to be an annotation file for topGO.
	GOAnnotations<-cbind(GOAnnotations$Group.1, GOAnnotations$x)
	row.names(GOAnnotations)<-GOAnnotations[,1]
	GOAnnotations<-GOAnnotations[,2]
	return(GOAnnotations)
}

#User function for conducting GO enrichment analysis of significantly upregulated and downregulated genes based on probe expression, using the csv file generated from ExtremesOfExpression.
```

#Can specify GO analysis of Biological Process "BP", Molecular Function "MF or Cellular Component "CC", can specify how many GO terms to list using ResultsToList parameter.

#Generates two CSV files of results, one for 'upregulated in high' and one for 'downregulated in high'.

DoGOAnalysis<-**function**(GOAnnotations, TopTableFile, GO_category=c("BP", "MF", "CC"), ResultsToList){

require(topGO)

options(width=600)

interestingGenes<-read.csv(TopTableFile, header=TRUE) #Load the csv file of the top 1000 (or those which adj.p.val<0.05, whichever is smaller) differentially expressed genes produced by ExtremesOfExpression.

i <- sapply(interestingGenes, is.factor)

interestingGenes[i] <- lapply(interestingGenes[i], as.character)

DOWN.in.HIGH<-unique(interestingGenes[interestingGenes$logFC>0, 9]) #Choose only those genes which are downregulated in high expressers (positive logFC), get a list of their gene symbols.

in.Genes.DOWN<-factor(as.integer(names(GOAnnotations) %in% DOWN.in.HIGH)) #Make a list of factors denoting which of genes in our universe are part of our interesting set of genes.

names(in.Genes.DOWN)<-names(GOAnnotations) #These factors must be labelled by gene name, in this case.

UP.in.HIGH<-unique(interestingGenes[interestingGenes$logFC<0, 9]) #Do the same for those genes upregulated in high expressers.

in.Genes.UP<-factor(as.integer(names(GOAnnotations) %in% UP.in.HIGH))

names(in.Genes.UP)<-names(GOAnnotations)

#Instantiate objects of class topGOdata to put our own data into.

```
GOdata_down<-new("topGOdata", ontology=GO_category, allGenes=in.Genes.DOWN, annot=annFUN.gene2GO, gene2GO=GOAnnotations, nodeSize=5)

GOdata_up<-new("topGOdata", ontology=GO_category, allGenes=in.Genes.UP, annot=annFUN.gene2GO, gene2GO=GOAnnotations, nodeSize=5)


#Conduct Fisher's Exact Tests to determine significantly enriched GO nodes, and store two objects containing the results

results.fisher.down<-runTest(GOdata_down, algorithm="classic", statistic="fisher")

results.fisher.up<-runTest(GOdata_up, algorithm="classic", statistic="fisher")


#Summarise the results in a table using the GenTable function.

top.x.results.down<-GenTable(GOdata_down, classicFisher=results.fisher.down, orderBy="classicFisher", topNodes=ResultsToList, numChar=300)

top.x.results.up<-GenTable(GOdata_up, classicFisher=results.fisher.up, orderBy="classicFisher", topNodes=ResultsToList, numChar=300)


Results<-NULL       #Create an object in which to store the results from both analyses.

Results$Up<-top.x.results.up

Results$Down<-top.x.results.down


names(Results$Up)<-c("GO ID", "Term", "Annotated", "Significant", "Expected", "p-Value") #Rename our table columns for presentation's sake.

names(Results$Down)<-c("GO ID", "Term", "Annotated", "Significant", "Expected", "p-Value") #Do the same with Down in High.
```

```
        write.csv(Results$Down, file=paste("DownInHigh.", ResultsToList, ".",
GO_category, ".", TopTableFile, sep=""))

        write.csv(Results$Up, file=paste("UpInHigh.", ResultsToList, ".",
GO_category, ".", TopTableFile, sep=""))

}

#####################################################################
```

APPENDIX 2: NEGATIVE CONTROL STAINING WITHOUT PRIMARY ANTIBODIES ON MOUSE TMA SECTIONS

		Mouse No Primary (Negative Control)					
		Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
Lung	1	X	C	C	C	C	C
	2	C	X	-	C	C	C
	3	C	C	C		C	C
	4	X	X	X	C	C	X
	5	C	X	X	C		C
	6	X	C	C	C	C	C
Kidney	1	X	C	C	C	C	C
	2	X	C	C	-	-	C
	3	-	-	X	C	C	X
	4	-	C	C	-	-	C
	5	X	X	X			X
Liver	4	-	X	X	C		X
	5	-	-	-	-	-	-
	6	X	-	X	-	-	-
	7	-	-	-	-	-	X
	8	-	-	-	-	-	Junctional
	9	X	-	Junctional	Junctional	Junctional	Junctional
	10	Junctional	X	-			X
Brain	1	-	-	-	-	-	-

	2	-	X	-	-	-	-
	3	-	-	C		-	-
Pancreas	1	-	C	C	-	-	-
	2	-	-	C	-	-	-
	3	-	-	X	-	-	X
Intestine	1	-	-	-	-	-	C
	2	-	-	-	-	-	C
	3	X	X	X	-	-	X
Skin	1	-	X	X	-	-	-
	2	-	X	X	-	-	C
	3	X	-	-			X
Spleen	1	C	-	X			X
	2	X	C	C	C	C	C
Skeletal Muscle	1	X	-	C	-	-	-
	2	-	-	X	C		X
Heart	1	C	-	C	-	C	C
	2	C	C	C	C	C	C
Ovary	1	-	X	X			X
	2	X	X	X			X
Testis	1	-	-	-	-	-	-
	2	-	-	-	-	-	-
Prostate	1	X	-	-	-	-	-
	2	X	-	-	-	-	-
	3	X	X	X	-	-	-
	4	-	-	-	-	-	-

5	X	-	-	-	-	-
6	X	X	X	-		X
7	-	-	-	-	-	-
8	-	-	-	-	-	-
% Core Loss	37.5	29.16666667	33.33333333	16.66666667	20.83333333	29.16666667

= absent core material

N = nuclear staining

C = cytoplasmic staining

M = membranous staining

Moderate blood vessel and extracellular staining is ignored.

		Rabbit No Primary Negative Control				
		Run 1	Run 2	Run 3	Run 4	Run 5
Lung	1	C	C	C	X	C
	2	X	X	C	C	X
	3	-	C	C	C	X
	4	X	X		X	X
	5	X	X	-	C	-
	6	C	C	C	C	-
Kidney	1	-	-	C	X	-
	2	-	X	C	C	-
	3	-	-	C	C	X
	4	-	-	-	-	-
	5	X	X		X	X
Liver	4	X	X		X	X
	5	-	-	-	-	-
	6	-	-	-	-	-
	7	-	-	-	-	-
	8	-	-	-	X	-
	9	-	X	-	-	-
	10	X	X		X	X
Brain	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
Pancreas	1	-	-	-	-	-
	2	-	X	-	-	-
	3	X	X	-	-	X

Intestine	1	-	X	-	-	-
	2	-	-	-	X	-
	3	X	-	-	-	X
Skin	1	-	-		X	X
	2	-	-		X	X
	3	-	X		X	X
Spleen	1	X	X		X	X
	2	-	-	C	C	-
Skeletal Muscle	1	-	-	-	C	X
	2	-	X		X	X
Heart	1	-	X	C	C	C
	2	C	C	C	X	C
Ovary	1.1	X	X		X	X
	1.2	X	X		X	X
Testis	1.1	-	-	-	-	-
	1.2	-	-	-	-	-
Prostate	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	X	-	X	-
	4	-	-	-	X	-
	5	-	-	-	-	-
	6	X	X		X	X
	7	X	X	-	C	-
	8	-	-	-	-	-
% Core Loss		27.08333333	41.66666667	25	39.58333333	37.5

 = absent core material

N = nuclear staining

C = cytoplasmic staining

M = membranous staining

Moderate blood vessel and extracellular staining is ignored.

APPENDIX 3: SUMMARY OF ANTI-ALDH7A1 IHC EXPERIMENTS ON MOUSE TMA SECTIONS

		ALDH7A1 1:200, Tris/EDTA Antigen Retrieval		
		Run 1	Run 2	Run 3
Lung	1	C	N	C
	2	X	X	C
	3	C	N/C	C
	4	X	X	C
	5	X	X	C
	6	C	N	C
Kidney	1	N	N	N
	2	N	X	N
	3	N	N	N
	4	N	N	N
	5	X	N	X
Liver	4	X	X	X
	5	N	N	N
	6	N	N	X
	7	N	N	N
	8	N	N	X
	9	X	X	N
	10	X	X	N
Brain	1	N	N	N
	2	N	N	N
	3	N	N	N
Pancreas	1	N	N	N

	2	N	N	N
	3	X	X	N
Intestine	1	-	X	C/N
	2	C/N	C/N	C/N
	3	X	C/N	C/N
Skin	1	C	C	-
	2	C	N	-
	3	C	X	C
Spleen	1	X	X	X
	2	C	C	C
Skeletal Muscle	1	C	C	C
	2	C	C	C
Heart	1	C/N	X	C/N
	2	C	C	N
Ovary	1.1	X	X	X
	1.2	X	X	X
Testis	1.1	C/N	C/N	C/N
	1.2	C/N	C/N	C/N
Prostate	1	-	C	C
	2	C	C	C
	3	X	X	C
	4	C	C	C
	5	C	C	C
	6	X	X	X
	7	X	X	C
	8	C	C	C

% Core Loss	31.25	35.4166667	16.6666667

X = absent core material

N = nuclear staining

C = cytoplasmic staining

M = membranous staining

- = no staining

Moderate blood vessel and extracellular staining is ignored.

APPENDIX 4: SUMMARY OF ANTI-ZSCAN4 IHC EXPERIMENTS ON MOUSE TMA SECTIONS

		ZSCAN4 1:200, Citrate Antigen Retrieval		
		Run 1	Run 2	Run 3
Lung	1	X	C	C
	2	C	X	C
	3	C	C	C
	4	X	X	C
	5	C	X	C
	6	X	C	C
Kidney	1	X	C	C
	2	X	X	C
	3	-	C	C
	4	-	C	C
	5	X	C	C
Liver	4	C	X	X
	5	-	C	C
	6	X	C	C
	7	C	C	C
	8	C	C	X
	9	X	X	C
	10	Junctional	X	X
Brain	1	-	-	C
	2	-	C	C
	3	-	C	C
Pancreas	1	C	-	-

	2	-	-	C
	3	-	X	-
Intestine	1	-	X	C
	2	-	C	C
	3	X	C	C
Skin	1	X	C	C
	2	C	C	C
	3	X	X	C
Spleen	1	C	X	X
	2	X	C	C
Skeletal Muscle	1	X	X	C
	2	C	C	C
Heart	1	C	X	C
	2	C	C	C
Ovary	1.1	C	X	X
	1.2	X	X	X
Testis	1.1	C	N/C	C
	1.2	C	C	C
Prostate	1	X	-	C
	2	X	C	C
	3	X	X	-
	4	-	C	C/M
	5	X	X	C
	6	X	X	X
	7	-	-	-
	8	-	C	C

% Core Loss	39.5833333	37.5	12.5

X = absent core material

N = nuclear staining

C = cytoplasmic staining

M = membranous staining

- = no staining

Moderate blood vessel and extracellular staining is ignored.

APPENDIX 5: BATH COHORT STUDY PROTOCOL AS APPROVED BY NHS R&D

Introduction

Prostate cancer is the most common male cancer in the UK and represents 25% of all new cancer diagnoses in males (Cancer Research UK). However, prostate cancer is a very heterogeneous disease with a large variation in clinical outcome. There is a pressing need to identify biological features (termed biomarkers) that can distinguish between aggressive tumours requiring radical intervention and those that have a good prognosis, where patients may be spared the potential side effects of treatment. The discovery of cancer stem cells is one of the most exciting recent advances in cancer research. Cancer stem cells are thought to have the ability to develop into all cancer cell types found in a tumour and may persist in tumours as a distinct cell population, offering an explanation for the post-treatment recurrence of cancers. It is hoped that understanding their behaviour will lead to the more effective diagnosis and treatment of cancer.

To this end, efforts are underway to find biomarkers that are suitable for identifying prostate cancer stem cells. We have collated published evidence from multiple publications using many experimental approaches - including cell line studies *in vitro*, xenograft transplantation studies *in vivo* and immunohistochemical examination of human tissue samples – in order to identify candidate biomarkers that may be associated with cancer stem cells in prostate cancer. Expression of candidate biomarkers will be examined in surplus prostate cancer specimens taken prior to primary treatment, in order to retrospectively identify potential predictive biomarkers of relapse. Such biomarkers would be useful for making initial treatment decisions in prostate cancer patients, and identify individuals at greater risk of relapse where more monitoring following treatment may be required. Additionally, if a group at low risk of relapse could be identified then it could help spare some patients the side effects of unnecessary treatment.

Background on Immunohistochemistry

Immunohistochemistry (IHC) is a method for detecting antigen (often protein) localisation (Ramos-Vara, 2005). Formalin-fixed and paraffin embedded tissue sections must be de-waxed and rehydrated in graded alcohols. To restore some of the

antigenicity that is masked during formalin fixation, antigen retrieval methods may then be used prior to staining, which consist of either heat or enzyme-mediated cleavage of methylene cross-links. The staining procedure involves primary antibodies specific to an antigen being adsorbed onto tissue sections 4-5um thick. Secondary antibodies which are conjugated to an enzyme are then applied to the section, causing a signal when substrates for this enzyme are added. In this way, coloured products are deposited at the sites of antigen localisation, allowing for semi-quantitative assessment of antigen abundance under a light microscope. The protocol for immunohistochemistry must be optimised separately for each antibody to be used. This requires that different combinations of antibody concentration and antigen retrieval techniques are trialled to establish the protocol which provides optimum signal strength and minimum non-specific staining. IHC is already well established in pathology laboratories and many antibodies are now used to assist in the diagnosis and staging of cancer.

Background on Tissue Microarrays

A tissue microarray (TMA) consists of an array of multiple paraffin-embedded tissue cores taken from formalin-fixed, paraffin-embedded tissue blocks (Camp *et al.*, 2008), such as those archived by hospital pathology departments. TMAs can be routinely sectioned, allowing many different tissues to be stained simultaneously on the same microscope slide. This conserves staining reagents and archived patient tissue while allowing for high-throughput analysis of many cases simultaneously. These properties make tissue microarrays ideal for biomarker studies.

Due to the small core sizes used in TMAs, there are concerns that staining a single core may be insufficient to obtain a representative sample in highly heterogeneous tumours such as breast and prostate cancers (Rubin *et al.*, 2002). There is some controversy over the size and number of cores that would be required to obtain a representative sample. According to Rubin et al., accurate estimates of the abundance of prostate cancer biomarkers - such as Ki-67 - may be better achieved when 3 or more 0.6mm cores are used. However, a recent study of prostate TMAs indicates that using multiple cores per tumour may introduce a bias towards marker positivity, suggesting that a single 0.6mm core is sufficient (Tennstedt *et al.*, 2012). In cases where the expression of markers is rare or focal, even large numbers of cores may be insufficient to assess protein abundance and result in false negative errors (Eckel-Passow *et al.*, 2010). To gather as much data as possible, two 1mm cores per tumour will be used for identifying candidate

biomarkers in this preliminary study. These have a total surface area equivalent to over five 0.6mm cores (Pinder *et al.*, 2013) and should provide a better representation of tumour heterogeneity. This sampling strategy has previously been employed in a large expression study of a candidate marker, BMI1(Sanchez-Beato *et al.*, 2006). This will be sufficient to evaluate if the candidate markers can be assessed using tissue microarrays. In subsequent studies the number and size of tissue cores may be reconsidered, and will depend on the abundance and heterogeneity of the markers to be studied.

Study Aims

This is a scientific research project to determine if markers for prostate cancer stem cells can be used to predict relapse in retrospective prostate cancer cases. Archived formalin fixed paraffin-embedded primary prostate cancer tissue from patients who have been treated with either hormone therapy or radical prostatectomy will be used in this study. Patients who have had neo-adjuvant hormone therapy followed by radiotherapy will be pooled with the hormone therapy patients. The aim is to identify markers whose expression is associated with relapse for each treatment group. Any markers identified in this initial study will be followed up in a subsequent study with a larger sample size to establish if they are useful biomarkers.

Study Design and Methods

Tissue microarrays will be constructed from archived formalin fixed paraffin-embedded prostate cancer tissue. Two cores of tumour tissue measuring 1mm in diameter will be taken from each tissue block, and a tissue microarray block will be prepared in order to immunohistochemically stain all sampled tissues simultaneously. This leaves the remainder of the tissue block intact if required clinically. Placement of cores will be guided by haematoxylin and eosin (H&E) stained whole sections that had previously been used for diagnosis. Cell lines with known expression of candidate markers will be fixed in formalin, embedded in agarose, processed and embedded in paraffin to form “cell blocks”. These cell blocks will be cored simultaneously with the tissues. These will serve as positive and negative controls in the tissue microarray, provided that these blocks are shown to be reliable. Alternatively, surplus archived human control tissue (for which informed consent will be obtained) can be used, where tissues with known expression of candidate markers will be cored simultaneously with the prostate samples and will serve as positive and negative controls in the tissue microarray. These tissues may be obtained by biopsy or autopsy. Control tissue may also be obtained from

recognised commercial suppliers which have obtained suitable informed consent. Control tissues that may be analysed (in addition to prostate cancer) include liver, breast, stomach, melanoma, pancreas, bone marrow or tonsil. Tissue microarray sections will be stained for the presence of putative cancer stem cell markers, such as BMI1 and Nestin. To confirm that the stained regions consistently represent cancer tissue, a single TMA section will be stained with routine H&E staining for every 25 tissue microarray sections examined – a greater level of quality control than the minimum suggested by Kumar *et al.*(2004). IHC stained sections will be examined microscopically using a semi-quantitative scoring system to assess the intensity and number of cells stained. Examples of scoring methods include the Allred scoring system (Allred *et al.*, 1998) – which adds scores for staining intensity and proportion of total cells stained to obtain a final score; the H-score system (Ishibashi *et al.*, 2003), where the percentage of cells staining in each of three intensity levels is multiplied by its intensity score and then totalled to give a range of values from 0-300; and automated scoring systems that make use of image analysis software (Hatanaka *et al.*, 2003). Scoring will be performed blinded, by multiple assessors if possible. For each treatment group, scores will be compared between those patients that have relapsed and those that remain in remission.

Primary Endpoint:

1. A difference between expression of candidate biomarkers in patients who have relapsed and patients who remain in remission after 5 years of retrospective follow-up. In both treatment groups, relapse is defined as serum PSA levels which persist at or rise above nadir and must be confirmed by a subsequent clinical evaluation.

There are no secondary endpoints for this study.

Study Monitoring and Data Handling

The research will be monitored internally by the University of Bath as part of the assessment of Ben Sharpe's PhD progress. He will submit a report after 11 months and have a formal viva to establish that sufficient progress is being made. He will be assessed at 24 months and finally by writing a thesis and defending it at a viva at the end of his studies (36-42 months).

Non-identifiable patient data will be provided in linked anonymised form. Individual patients will be referred to by randomly generated code numbers. Only authorised

personnel at RUH will have access to the personal data. No identifiable data will be provided to researchers. Data will be collected and retained in accordance with the Data Protection Act 1998. Data will be retained in a secure location during and after the trial has finished (for 10 years after the end date of the study). Any tissue samples remaining at the end of the study will be returned to the Department of Pathology, RUH.

Inclusion Criteria

Inclusion criteria must be met at the time of entry into the study.

- Age ≥ 18 years
- Diagnosis of prostate adenocarcinoma at time of sampling. This must be histologically confirmed and the original haematoxylin & eosin stained slides must be available for the case.
- Patients treated with radical prostatectomy must be stages T1-3 N0 M0. Patients treated with hormone therapy must be stages T1-3 N0-1 M0-1.
- Signed informed consent must be obtained prior to any archived tissue samples being taken.
- Patient must have undergone radical prostatectomy. Alternatively, patients must have undergone core needle biopsy prior to hormone therapy.
- All patients must have archived primary tumour tissue in formalin-fixed and paraffin-embedded form. This tissue must be surplus to clinical requirements.
- For patients treated with Radical Prostatectomy, patients must have no evidence of metastatic disease at the time of sample acquisition.
- Minimum 5 years of retrospective follow-up time for patients in the non-relapse category.
- Patients with stored non-identifiable clinical data.

Exclusion Criteria

Patients who meet the following criteria during the study duration will be excluded:

- Previous or concurrent cancer (at time of treatment) that is distinct in primary site or histology from the cancer being evaluated in this study EXCEPT treated basal cell carcinoma, and superficial bladder tumours [Stage Ta, Tis & T1].
- Patients that have undergone radical prostatectomy with positive lymph node involvement [Stage N1] or metastases [Stage M1].

- All patients that have not been assessed for lymph node involvement [Stage NX] or metastases [Stage MX].
- Patients treated with neoadjuvant hormone therapy or chemotherapy prior to radical prostatectomy.
- Insufficient amounts of archived tumour tissue available from the patient. Tissue blocks must be at least 1mm thick in order to provide an adequate core depth for microarray construction.
- Withdrawal of consent for the use of archived patient tissue.

Statistical Considerations

The aim of this study is to identify potential predictive biomarkers of relapse in prostate cancer. The study will look at markers in patients following two treatment types. The first treatment type will be radical prostatectomy, and 20 samples from patients who relapsed and 20 samples from patients who remain in remission will be analysed for each marker. The second treatment type will be hormone therapy (either as a primary therapy or as neoadjuvant therapy prior to radiotherapy) and 20 samples from patients who relapsed and 20 samples from patients who remain in remission will be analysed for each marker. The two groups will be non-overlapping so there will be a total of 80 samples analysed.

Power calculations carried out by Dr Taylor (Reader in Medical Statistics, Department for Health, University of Bath) indicated that 13 patients per group would be sufficient to identify a difference. The calculation assumed that a marker showing a difference would be scored positive in 80% of cases with relapse and 20% of cases without relapse. Our aim of 20 samples per group exceeds this minimum value.

Commonly used scoring systems, such as the Allred system (Allred *et al.*, 1998), will be used to analyse the IHC staining. The Allred system gives one score for intensity of stain and one for the proportion of cells stained (%) and these are combined to create a score out of eight for each marker. If a tumour receives a score of 3 or more it is considered positive for that marker. For each primary tumour, each marker will be classified as positive or negative and the data will be collated and statistically analysed as appropriate. Confidence intervals will be calculated to estimate the sensitivity and specificity of the marker. Any marker which shows a correlation with tumour relapse may have the potential to be used as a biomarker and would warrant further analysis in subsequent studies.

Ethical Considerations

This study involves immunohistochemical analysis of archived patient prostate specimens. Written informed consent will be obtained from each patient before the study begins. Tissue used will be archived formalin fixed paraffin-embedded tissue which is surplus to clinical requirement. Archived patient tissue and non-identifiable data will be in “linked anonymised form” and referred to by a randomly generated code number. This tissue will be securely stored in the Department of Pathology at the RUH and in the Department of Biology and Biochemistry at the University of Bath. On completion of the study all tissue will be returned to the site of origin – the Department of Pathology, RUH. Signed informed consent can be withdrawn at any time during the study, at which point the withdrawn tissue and corresponding anonymised data will no longer be used. No identifiable data will be released to researchers.

Ethics and R&D approval

This study is subject to Research Ethics Committee (REC) and local Research and Development (R&D) approval.

APPENDIX 6: INVITATION LETTER FOR BATH COHORT STUDY

RUH Royal United Hospital Bath NHS
NHS Trust

RUH Cancer Clinical Trials Unit

Royal United Hospital,
Combe Park,
Bath,
BA1 3NG
Tel: 01225 825096

Prostate Cancer Biomarker Study Invitation Letter

Dear

We would like to invite you to take part in our research study which is looking at new laboratory tests for prostate cancer. You are being invited as you have been treated for prostate cancer at the RUH.

If you decide to take part you will need to discuss the project with us and give us written consent to study some of your existing tissue samples. You would not require any new procedures or tests and it would not affect your future treatment or follow-up.

If you would be interested in learning more about this study and possibly participating, please let us know by returning this letter in the envelope provided and we will be in touch.

Thank you for your time in reading this letter.

With kind regards,

Dr Mark Beresford and Dr Rebecca Bowen
Consultant Oncologists

I would be happy for you to contact me with regard to participating in the study ☐

Preferred contact telephone number: _____

Preferred day and time for us to contact you: _____

APPENDIX 7: BATH COHORT PATIENT INFORMATION SHEETS

RUH Royal United Hospital Bath NHS
NHS Trust

RUH Cancer Clinical Trials Unit

Tel: 01225 825096

Patient Information Sheet
Prostate Cancer Biomarker Study

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it would involve for you. **One of our team will go through the information sheet with you and answer any questions you have**. You may find it helpful to talk to others about this study. Before you decide whether or not to take part it is important you understand why this research is being done and exactly what is involved. Please read the following leaflet carefully and ask the nurse or doctor if there is anything that is unclear or if you require any further information.

Part 1 tells you the purpose of this study and what will happen to you if you take part.

Part 2 gives you more detailed information about the conduct of the study.

Part 1

What is the purpose of this study?
Prostate cancer is the most common male cancer in the UK. Many men with prostate cancer have a very good prognosis and could potentially be spared the side effects of treatment. There is, however, a pressing need to identify biological features (termed biomarkers) that can distinguish the more aggressive prostate cancers requiring more radical treatment.
Cancer stem cells are thought to have the ability to develop into any cancer cell type and offer an explanation for the recurrence of cancer after the completion of treatment. It is hoped that understanding their behaviour will lead to more effective and individualised treatment of cancer.

What does the study aim to do?
This project will examine if expression of stem cell markers might be used to identify the aggressive variants of prostate cancer. To do this we will use immunohistochemistry tests, a well-established approach, in prostate cancer samples. In order to have been told you have a diagnosis of prostate cancer you will have had either a prostate biopsy or a prostatectomy (removal of the prostate gland). After initial diagnostic tests are complete, these tumour samples are stored
Patient Information: Prostate Cancer Biomarker Study Version 2 (25/6/2013)

as blocks in the Pathology department of the hospital. It is these stored blocks, now surplus to clinical requirement that we would like to use to examine the expression of stem cell markers. These markers, or subset of them, may enable us to identify the aggressive relapsing prostate cancers.

Why have I been invited?
You have been invited because you are currently being treated or were previously treated or for prostate cancer at the RUH. The study aims to recruit 80 participants.
Do I have to take part?
No. It is entirely your choice. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason, without affecting your care in any way.

What will happen to me if I take part?
If you agree to join the study then some of your prostate tissue which is currently stored at the RUH will be used. The stored tissue was originally taken and used to help with your diagnosis. A small part of this stored tissue, which is not needed for your treatment, will be used. **No new tissue samples will be taken.** You will not be asked to undergo any extra tests or procedures in order to participate in this study.

Expenses and payments
We are unable to offer payment for participation in this research but you should not incur any extra expenses as a result of participation.

What will I have to do?
Should you agree to participate we will ask you to sign a consent form. You will not be required to do anything else.

What are the possible advantages and disadvantages of taking part?
As this study does not involve any new procedures or treatments there should no risk or disadvantage in taking part. This study may not benefit you directly but the information we get from this study may impact on the future treatment of prostate cancer.

What if there is a problem?
Any complaint about the way you have been dealt with during the study will be addressed. The detailed information on this is given in Part 2.

Will my taking part in the study be kept confidential?
Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before you make any decision.

Part 2

What will happen if I don't want to carry on with the study?
If you decide to withdraw from the study then information collected using your tissue samples will be removed from the study and destroyed. No further information will be collected. You are free to withdraw consent at any time and this will not adversely affect your treatment.

What if there is a problem?
If you have a concern about any aspect of this study, the researchers will do their best to answer your question (see phone number above). If you remain unhappy and wish to complain formally, you can do this through the Patient Advice and Liaison Service (PALS).

Email: ruh-tr.PatientAdviceandLiaisonService@nhs.net
Tel: 01225 825656
Address: Head of PALS, RUH Bath NHS Trust
Combe Park
Bath
BA1 3NG

Will my taking part in this study be kept confidential?
All information which is collected about you during the course of the research will be kept strictly confidential, and any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised.

What will happen to any samples I give?
A small part of your existing sample which was used to assist in your diagnosis, currently stored at the RUH, will be used. Samples used will be stored securely at the RUH or at the University of Bath, where some of the analysis will take place. Only members of the study team will have access to the samples and it will not be possible to identify you from these samples. At the end of the study the samples will be returned to the RUH or destroyed.

Will any genetic tests be done?
No genetic tests will be carried out.

What will happen to the results of the research?
The nature of the research is such that it is very unlikely to have any direct impact on you or your management. Results will not be relayed directly back to you, your hospital doctor or GP. The results of the project, which may not be available for several years, may be published in scientific journals or conferences in anonymised form so you cannot be identified.

Who is organising and funding the research?
The research is being organised by clinicians from the RUH (Dr Mark Beresford, Dr Rebecca Bowen and Dr John Mitchard) and researchers from the University of Bath (Dr Andrew Chalmers and Mr Ben Sharpe). The project is being funded by

the University of Bath and the Annet Trust, which is a local charity which funds medical research projects.

Who has reviewed the study?
All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by The Proportionate Review Sub-Committee of the West of Scotland Research Ethics Committee 5.

What if I have further questions?

If you have any further queries about tissue donation or about research projects is general please contact the Clinical Trials Unit at the Royal United Hospital, Bath. Tel: **01225 825096.** For advice on whether to participate please ask a member of the study team, who can be contacted via the clinical trials unit (Tel: 01225 825096), or one of your other health care professionals.

Many thanks for taking the time to read this information sheet and for considering being involved in our research.

Dr Mark Beresford and Dr Rebecca Bowen
Consultant Oncologists
Royal United Hospital Bath

APPENDIX 8: BATH COHORT STUDY CONSENT FORM

Centre Number:
Study Number:
Patient Identification Number for this trial:

Royal United Hospital Bath NHS
NHS Trust

CONSENT FORM

Title of Project: **Identifying new biomarkers for prostate cancer**

Name of Researcher: Dr Mark Beresford and Dr Rebecca Bowen

Please initial box

1. I confirm that I have read and understand the information sheet dated..... (version.....) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that the donation of this tissue is voluntary and that I am free to withdraw consent at any time without giving any reason, without my medical care or legal rights being affected. ☐

3. My donated tissue and medical details will be anonymised. Relevant information from my medical records will be recorded and held on a computer database and supplied anonymously to the researcher. ☐

4. Access to personal identifiable data will be restricted to approved personnel only ☐

5. Any tissue which remains unused after 30 years will be disposed of in a lawful manner ☐

6. I understand the tissue used in this research is surplus to clinical requirements and would otherwise be discarded. ☐

7. I agree to my GP being informed of my participation in the study. ☐

8. I agree to take part in the above study. ☐

Name of Patient .....

Signature ..... Date.....

Name of Person taking consent .....

Signature ..... Date.....

Consent form date of issue: Consent form version number: 1 (10/513)

When completed: 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

Chairman, Brian Stables
Chief Executive, James Scott

Healthcare you can Trust